

Modellierung navigatorischer Fahrerpräferenzen
für die personalisierte Routenberechnung

MODELLIERUNG NAVIGATORISCHER FAHRERPRÄFERENZEN FÜR DIE PERSONALISIERTE ROUTENBERECHNUNG

Von der
Carl-Friedrich-Gauß-Fakultät
der Technischen Universität Carolo Wilhelmina zu Braunschweig

zur Erlangung des Grades eines
Doktors der Wirtschaftswissenschaften (Dr. rer. pol.)
genehmigte Dissertation

von
Andreas Landau
geboren am 16.10.1980
in Clausthal-Zellerfeld

Eingereicht am: 08.02.2016

Disputation am: 22.04.2016

1. Referent: Prof. Dr. Dirk Christian Mattfeld

2. Referent: Prof. Dr. Jan-Fabian Ehmke

3. Referent: Prof. Dr. Mark Vollrath

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

1. Aufl. - Göttingen: Cuvillier, 2016

Zugl.: (TU) Braunschweig, Univ., Diss., 2016

© CUVILLIER VERLAG, Göttingen 2016

Nonnenstieg 8, 37075 Göttingen

Telefon: 0551-54724-0

Telefax: 0551-54724-21

www.cuvillier.de

Alle Rechte vorbehalten. Ohne ausdrückliche Genehmigung des Verlages ist es nicht gestattet, das Buch oder Teile daraus auf fotomechanischem Weg (Fotokopie, Mikrokopie) zu vervielfältigen.

1. Auflage, 2016

Gedruckt auf umweltfreundlichem, säurefreiem Papier aus nachhaltiger Forstwirtschaft

ISBN 978-3-7369-9280-1

eISBN 978-3-7369-8280-2

Für meine Eltern

und

für meine Frau Katja

Vorwort

Die vorliegende Arbeit entstand während meines Promotionsstudiums an der Technischen Universität Braunschweig und wurde von der Carl-Friedrich-Gauß Fakultät im April 2016 als Dissertation angenommen.

Ein besonders herzlicher Dank gilt meinem Doktorvater, Herrn Prof. Dr. Dirk C. Mattfeld, für die vertrauensvolle und sehr konstruktive Zusammenarbeit. Er war stets an neuen Ergebnissen interessiert und unterstützte mich mit wertvollen Hinweisen und regelmäßiger Gesprächsbereitschaft. Ein sehr großer Dank geht auch an Herrn Prof. Dr. Jan-Fabian Ehmke, der mich sowohl als Assistent am Institut wie auch als Professor an der Freien Universität Berlin thematisch begleitet und wertvolle Anregungen gegeben hat. Ebenfalls danke ich Herrn Prof. Dr. Mark Vollrath für die Arbeit als Drittgutachter sowie Herrn Prof. Dr. Marc Gürtler für die Übernahme der Aufgabe des Prüfungsvorsitzenden.

Dazu möchte ich mich bei Herrn Ulrich Kersken und Herrn Guido Müller von der Robert Bosch Car Multimedia GmbH bedanken, die das Thema angestoßen und begleitet haben. Ebenso bei Herrn Prof. Dr. Sándor Fekete für wichtige grundlegende Hinweise und Anmerkungen für die Arbeit. Ganz herzlich bedanke ich mich bei Frau Dr. Barbara Metz von der WIVW GmbH sowie Frau Mariana Just von der BMW AG und Herrn Friedemann Kuhn von der Daimler AG für die Unterstützung sowie die Bereitstellung der Daten für die empirischen Analysen.

Meinen Freunden und meiner Familie danke ich für die positive Motivation und das Interesse an meiner Arbeit. Ein ganz besonderer Dank gilt auch meiner Frau Katja, die mich immer wieder motiviert hat, diese Arbeit zu erstellen und mir mit Rat und Tat zur Seite stand.

Veitshöchheim im Juni 2016

Andreas Landau

Zusammenfassung

Die Routingkomponente heutiger Navigationssysteme basiert auf einem Kürzeste-Wege-Verfahren, das für eine Zielfunktion einen Weg auf einem Graphen mit Kantengewichten berechnet. Standardmäßig wird die Minimierung der Routendauer oder der Routenlänge als Zielfunktion für die Route genutzt, um die navigatorischen Präferenzen der Fahrer zu erfüllen.

In dieser Arbeit wird eine Personalisierung der Routingkomponente durchgeführt, um so Unterschiede zwischen den Fahrern zu berücksichtigen und ihre navigatorischen Präferenzen in das Kürzeste-Wege-Verfahren einzubeziehen. Die Personalisierung umfasst eine automatische Anpassung der Zielfunktion sowie der Kantengewichte an den Fahrer. Hierfür wird ein Modell der navigatorischen Präferenzen des Fahrers erstellt, das auf Analysen verschiedener Perspektiven des Navigationsverhaltens der Fahrer basiert. Die Präferenzen werden dabei sowohl bezüglich der Dimension der Zeit wie auch bezüglich der Dimension des Raumes modelliert.

Für die erste Perspektive werden die gemessenen Geschwindigkeiten auf den Kanten untersucht. Das entsprechende Modell dieser Kantenperspektive gehört zur Dimension der Zeit. Das zweite raumorientierte Modell wird aus der Analyse der Gesamtheit aller Kanten eines Fahrers generiert (Netzwerkperspektive). Für die dritte Perspektive werden die vom Fahrer gewählten Routen analysiert und zwei Modelle beider Dimensionen erstellt.

Die Modelle werden statistisch untersucht und zu einem zweidimensionalen Modell der navigatorischen Präferenzen des Fahrers aggregiert. Das Ergebnis ist eine Zuordnung der Fahrer zu einem Routingtyp. Auf Basis der Modellierung werden sowohl die Zielfunktion wie auch die Kantengewichte zunächst separat und anschließend kombiniert personalisiert. Im Vergleich zu einer Standardroute erfolgt eine Bewertung der Personalisierungsansätze mit Hilfe verschiedener Gütemaße. Das entwickelte Verfahren für die Modellierung und die Personalisierung wird auf einen umfangreichen Realfahrdatensatz angewendet und zeigt so den Nutzen einer modellgestützten Personalisierung der Routingkomponente auf.

Abstract

The routing component of nowadays navigation devices bases on a shortest path calculation using an objective function and edge attributes of an underlying graph. Minimizing the route duration or route length are the most common objective functions to meet drivers' general navigation preferences for a route.

In this work, we include the navigation preferences of different types of drivers in the path calculation by personalizing the routing component. The objective function as well as edge attributes are adapted to drivers' preferences in an automated way. To describe drivers' navigation preferences in a navigation context, it is necessary to model driver behavior appropriately. Three different models are developed by analyzing different perspectives of drivers' navigation behavior covering the dimension of time as well as the dimension of space.

The first, time-oriented model examines the observed speeds on the edges. For the second, space-oriented model, a measure considering the entirety of edges is provided as to characterize the coverage of a driver's mobility network. The third model analyses routes consisting of paths of edges considering both dimensions. Correlation analyses show that results obtained from the different models confirm each other. Based on the results, a well-founded two-dimensional model is built from well-known attributes and to classify the drivers in respect to their navigation preferences.

This model is used for personalizing the objective function and the edge attributes separately and in a combined way. The calculated routes are compared with a standard route in respect to different measures to evaluate the personalization. All steps of modelling and personalization are implemented for a large scale data set. In this way, the improvement of the navigation device by a model-based personalization is quantified for each driver as well as for the determined classes of the sample.

Inhaltsverzeichnis

Inhaltsverzeichnis	IX
Abbildungsverzeichnis	XIII
Tabellenverzeichnis	XVII
Abkürzungsverzeichnis	XVII
Variablenverzeichnis	XXI
1 Einleitung	1
2 Grundlagen der Navigation	7
2.1 Navigationssysteme	7
2.1.1 Entwicklung der Navigation	7
2.1.2 Markt für Navigationslösungen	11
2.2 Komponenten von Navigationssystemen	15
2.2.1 Digitale Karte	16
2.2.2 Ortung	19
2.2.3 Map Matching	21
2.2.4 Routing	23
2.2.5 Human-Machine-Interface	25
2.3 Verbesserung der Routingkomponente der Navigation	27
2.3.1 Ansatzpunkte zur Verbesserung	27
2.3.2 Einbeziehung externer Einflüsse	29
2.3.3 Einbeziehung des Fahrers	33
2.3.4 Adaptierung durch den Fahrer	34
2.3.5 Personalisierung durch Adaption an den Fahrer	36
2.4 Analyse existierender Personalisierungsansätze und -verfahren	38
2.4.1 Verfahren in nicht-fahrzeugbezogenen Gebieten	39
2.4.2 Implementierte Ansätze im Fahrzeug	42
2.4.3 Forschungsansätze für Navigationssysteme	44

2.4.4	Bewertung der Ansätze und Verfahren	49
3	Theoretischer Ansatz für eine modellgestützte Personalisierung	51
3.1	Personalisierung der Routingkomponente	52
3.1.1	Entwicklung des Konzepts	52
3.1.2	Methoden der Bewertung	54
3.2	Modellierung der navigatorischen Präferenzen	56
3.2.1	Anforderungen	56
3.2.2	Dimensionen und Perspektiven der Modellierung	57
3.3	Vorgehen in Anlehnung an einen Knowledge Discovery Process . . .	60
4	Datenbasis der empirischen Untersuchungen	65
4.1	Beschreibung der Datenbasis	65
4.1.1	Datenaufzeichnung	66
4.1.2	Kennzahlen der Datenbasis	68
4.2	Einordnung der Datenbasis	70
4.2.1	Fahrertypen	71
4.2.2	Mobilität in Deutschland 2008	72
5	Verfahren zur modellgestützten Personalisierung	73
5.1	Schritte des Verfahrens	73
5.1.1	Datenaufbereitung	75
5.1.2	Analyse und Modellierung	77
5.1.3	Personalisierung und Bewertung	78
5.2	Analyse der navigatorischen Präferenzen der Fahrer	79
5.2.1	Kantenperspektive	79
5.2.2	Routenperspektive	81
5.2.3	Netzwerkperspektive	84
5.3	Modellierung auf Basis der Ergebnisse der Perspektivenanalysen . .	90
5.3.1	Korrelationsanalyse	90
5.3.2	Erstellung eines zweidimensionalen Modells	91
5.4	Durchführung der Personalisierung	92
5.4.1	Ansatz der personalisierten Kantengewichte	94
5.4.2	Ansatz der personalisierten Zielfunktion	95
5.4.3	Kombinierter Personalisierungsansatz	96
5.5	Bewertung der Personalisierung	97
5.5.1	Gütemaß der Zielkriterien	98
5.5.2	Alternative Gütemaße	101

6 Datenaufbereitung	105
6.1 Datenauswahl	106
6.2 Datenbereinigung	109
6.2.1 Invalide Positionsdaten	110
6.2.2 Kartendaten	112
6.2.3 Sprünge valider Positionsdaten	114
6.2.4 Matching	116
6.3 Datenreduktion	119
6.3.1 Grundlage für die Kantenperspektive	119
6.3.2 Grundlage für die Routenperspektive	121
6.3.3 Grundlage für die Netzwerkperspektive	122
6.3.4 Digitale Karte für die interne Routingreferenz	125
7 Empirische Untersuchungen zur modellgestützten Personalisierung	127
7.1 Ergebnisse für die Kantenperspektive	127
7.1.1 Anpassung der Kantengewichte für die Personalisierung . . .	127
7.1.2 Erstellung des Modells	130
7.2 Ergebnisse für die Routenperspektive	132
7.3 Ergebnisse für die Netzwerkperspektive	135
7.3.1 Indikator der Größe	136
7.3.2 Indikator der Struktur	138
7.3.3 Bestimmung eines Navigationstyps für die Modellerstellung .	141
7.4 Erstellung eines zweidimensionalen Modells	143
7.4.1 Durchführung der Korrelationsanalyse	143
7.4.2 Ergebnis der zweidimensionalen Modellierung	146
7.5 Umsetzung der Personalisierung	148
7.5.1 Personalisierung der Kantengewichte	148
7.5.2 Personalisierung der Zielfunktion	150
7.5.3 Kombinierte Personalisierung	151
7.6 Ergebnisse der Bewertung der Personalisierung	152
7.6.1 Nutzung der Zielkriterien als Gütemaß	153
7.6.2 Nutzung alternativer Gütekriterien	159
7.6.3 Diskussion der Bewertungsergebnisse	165
8 Fazit und Ausblick	167
Literaturverzeichnis	173

Abbildungsverzeichnis

2.1	Absatzzahlen von festeingebauten Navigationssystemen	8
2.2	Absatzzahlen von Navigationssystemen	9
2.3	Entwicklungen im Bereich der Navigation	10
2.4	Absatzzahlen im Mobilfunkbereich	12
2.5	Umsatz im AM-Bereich	13
2.6	Absatzzahlen im globalen Markt für Navigation	13
2.7	Umsatzzahlen im globalen Markt für Navigation	14
2.8	Vereinfachter Aufbau von Navigationssystemen	15
2.9	Illustration eines gerichteten Graphen	16
2.10	Illustration eines Ausschnitts der digitalen Karte	17
2.11	Visualisierung von OpenStreetMap	17
2.12	Vereinfachter Aufbau der digitalen Karte	18
2.13	Schematische Darstellung der satellitengestützten Positionsbestimmung	19
2.14	Aufbau des Ortungsverfahrens	20
2.15	Illustration des Map Matchings	21
2.16	Ablauf des trassenbasierten Map Matchings	22
2.17	Illustration der Herausforderungen des Map Matching Verfahrens .	23
2.18	Illustration eines kürzesten Weges	24
2.19	Lösungsraum der Routenberechnung	24
2.20	Übersicht gängiger Kürzeste-Wege-Verfahren	25
2.21	Beispiel der Gestaltung der Route und des Menüs	26
2.22	Einflüsse auf den berechneten Zielfunktionswert der Routenberechnung	28
2.23	Visualisierung der Möglichkeiten zur Beeinflussung der Menge der zulässigen Routen sowie der optimalen Route	29
2.24	Visualisierung von Ganglinien	30
2.25	Vorgehen eines Verkehrsmeldungsdienstes	31
2.26	Grafische Veranschaulichung des aktuell herrschenden Verkehrs . . .	32
2.27	Routenauswahlverhalten des Fahrers und des Navigationssystems .	33

2.28	Adaptierung des Navigationssystems durch den Fahrer	36
2.29	Adaption des Navigationssystems an den Fahrer	37
2.30	Prozess der Personalisierung	38
2.31	Klassifizierungsprozess beim Collaborative Filtering	40
2.32	Zuordnungsprozess beim Collaborative Filtering	40
2.33	Schaubild des Systems Nutzer-Computer	41
2.34	Schaubild des Systems Fahrer-Fahrzeug-Umwelt	42
2.35	Schaubild des Systems Fahrer-Fahrzeug-Umwelt nach FAS	43
2.36	Darstellung eines adaptiven Bremsassistenten	43
2.37	Darstellung des 3-Ebenenmodells der Fahraufgabe	44
2.38	Routenauswahlkriterien	46
2.39	Einführung neuer Zielkriterien	48
3.1	Veranschaulichung der Gewichtung der Dimensionen Raum (fix) und Zeit (variabel) bei der Routenberechnung	53
3.2	Bewertung auf Basis der Übereinstimmung	54
3.3	Bewertung auf Basis der Straßenklasse	55
3.4	Bewertung auf Basis der Zielkriterien	56
3.5	Anforderungen an die Modellierung	57
3.6	Visualisierung der potentiellen Dimensions- und Perspektivenkorre- lationen	59
3.7	Deskriptive Beschreibung der Routingtypen	59
3.8	Überblick über die Datenverarbeitungsschritte eines KDPs	61
3.9	Aufbau des CRISP-DM KDPs	62
3.10	Struktur des Vorgehens für die Modellierung und Personalisierung der navigatorischen Präferenzen eines Fahrers	63
4.1	Standbild der Videoaufzeichnung im Fahrzeugtyp 1 (links) und Fahr- zeugtyp 2 (rechts).	67
4.2	Grafische Darstellung aller aufgezeichneten Fahrten	69
5.1	Aufbau des Verfahrens	74
5.2	Ablauf der Analyse für die Kantenperspektive	80
5.3	Struktur der externen Routingreferenz	82
5.4	Ablauf der Analyse für die Routenperspektive	83
5.5	Ablauf der Analyse für die Netzwerkperspektive	85
5.6	Illustration eines Kreises	86
5.7	Illustration der Kerne	87
5.8	Illustration der verschiedenen Navigationstypen	89

5.9	Ablauf der Korrelationsanalyse	91
5.10	Ablauf der zweidimensionalen Modellierung	92
5.11	Visualisierung der unterschiedlichen Personalisierungsansätze	93
5.12	Struktur der internen Routingreferenz	94
5.13	Ablauf der Personalisierung der Kantengewichte	95
5.14	Ablauf der Personalisierung der Zielfunktion	96
5.15	Ablauf der kombinierten Personalisierung	97
5.16	Vorgehen zur Bewertung auf Basis der Zielkriterien	99
5.17	Vorgehen zur Analyse der Perspektivenmodelleinflüsse	101
5.18	Vorgehen zur Bewertung auf Basis der Straßenklassenanteile	102
5.19	Vorgehen zur Bewertung auf Basis der Übereinstimmung	104
6.1	Ablauf der Datenaufbereitung	105
6.2	Ablauf der Datenauswahl	107
6.3	Ablauf der Datenbereinigung für eine Fahrt	110
6.4	Vorgehen zur Bereinigung der Positionsdaten	111
6.5	Vorgehen zur Bereinigung der Kartendaten	113
6.6	Beispiele für die Bereinigung der Kartendaten	114
6.7	Vorgehen zur Bereinigung von Sprüngen	115
6.8	Ablauf der Matchingverbesserung	116
6.9	Visualisierung der Matchingverbesserung	118
6.10	Visualisierung des Ergebnisses der Matchingverbesserung	118
6.11	Häufigkeitsverteilung des Benutzungsgrads der Kanten eines Fahrers	120
6.12	Anzahl der ausgewählten Kanten eines Fahrers	120
6.13	Visualisierung einer Route	121
6.14	Anzahl FR-Routen länger als 1 km pro Fahrer	122
6.15	Ablauf der Netzwerkerstellung	123
6.16	Visualisierung der Netzwerkerstellung	125
6.17	Ablauf der Kartenerstellung	126
6.18	Visualisierung der erstellten Karte	126
7.1	Anzahl Kanten pro Geschwindigkeitsklasse für jeden Fahrer	128
7.2	Gemessene Geschwindigkeit pro Geschwindigkeitsklasse für jeden Fahrer	129
7.3	Bewertung der gemessenen Geschwindigkeiten pro Fahrer	130
7.4	Unterschied zwischen theoretischer und gemessener Geschwindigkeit pro Geschwindigkeitsklasse für jeden Fahrer	131
7.5	Ergebnis der Modellierung der Kantenperspektive pro Fahrer	132

7.6	Anzahl berechneter Routen	133
7.7	Anteil berechneter Routen	134
7.8	Ergebnis der Modellierung der Routenperspektive pro Fahrer	135
7.9	Verteilung der Knoten- und Kantenanzahl sowie der Kantengewichte	136
7.10	Normierte Kennwerte für Knoten, Kanten und Kantengewichte . . .	137
7.11	Wert des Indikators Größe pro Fahrer	138
7.12	Verteilung der Anzahl an Kreisen sowie an Knoten in den Kernen .	139
7.13	Relativer Kennwert der Kerne	140
7.14	Normierte Kennwerte für Kreise sowie der relativen Kerngrößen . .	140
7.15	Wert des Indikators Struktur pro Fahrer	141
7.16	Visualisierung des Navigationstyps	142
7.17	Ergebnis der Modellierung der Netzwerkperspektive pro Fahrer . . .	143
7.18	Vergleich der Kennwerte gleicher Dimension	144
7.19	Vergleich der Kennwerte unterschiedlicher Dimension	145
7.20	Mittelwertbasierte Unterteilung der Fahrer in vier Klassen	146
7.21	Klassifizierung der Fahrer mit Hilfe eines Clusterverfahrens	147
7.22	Abweichung der gemessenen von denen in der digitalen Karte ge- speicherten Geschwindigkeiten	149
7.23	Abweichung der Geschwindigkeit für die verschiedenen Geschwin- digkeitsklassen, Ausschnitt	150
7.24	Anteile der zweidimensionalen Zielfunktion, sortierte Darstellung . .	151
7.25	Anzahl berechneter Routen pro Fahrer	152
7.26	Abweichung der Routendauer	153
7.27	Abweichung der Routenlänge	154
7.28	Bewertung der Routendauerabweichung	155
7.29	Bewertung der Routenlängenabweichung	156
7.30	Bewertung der aggregierten Zielkriterienabweichung	157
7.31	Einfluss der Perspektivenmodelle	159
7.32	Bewertung der berechneten Routen für die kantenanzahlbasierten Straßenklassenanteile	160
7.33	Bewertung der berechneten Routen für die längenbasierten Straßen- klassenanteile	161
7.34	Nutzen der Personalisierung für die kantenbasierten Straßenklassen- anteile	162
7.35	Nutzen der Personalisierung für die längenbasierten Straßenklassen- anteile	163
7.36	Bewertung der berechneten Routen für die Übereinstimmung	164
7.37	Nutzen der Personalisierung für die Übereinstimmung	165

Tabellenverzeichnis

3.1	Perspektiven und Dimensionen der Modellierung	58
4.1	Übersicht über die Datenaufzeichnung	66
4.2	Übersicht über die Datenbasis	68
4.3	Beschreibung der erhobenen Daten im Untersuchungszeitraum pro Fahrer	69
4.4	Stichprobenbeschreibung auf Basis von Fragebögen	70
4.5	Vergleich der euroFOT-Stichprobe mit der MiD2008-Studie	72
5.1	Unterteilung der Routen	76
5.2	Übersicht über die verschiedenen Navigationstypen	89
5.3	Übersicht über die für die Bewertung genutzten Routen	98
6.1	Ausgewählte Variablen der Datenbasis	108
6.2	Kennzahlen der Datenbasis nach der Datenauswahl	109
7.1	Korrelationen zwischen den Kennwerten der verschiedenen Perspek- tiven und Dimensionen	145

Abkürzungsverzeichnis

ACC	Adaptive Cruise Control
ADAC	Allgemeiner deutscher Automobilclub
AM	Aftermarket
CAN	Controller Area Network
CD	Compact Disk
CRISP-DM	Cross-Industry Standard for Data Mining
D	Deutschland
DBQ	Driver Behaviour Questionnaire
DGPS	Differential Global Positioning System
DR	Dead Reckoning
DVD	Digital Versatile Disc
E	Electronic
EU	Europäische Union
FAS	Fahrerassistenzsysteme
FR	Fahrer-Real
FOT	Field-Operational-Test
GLONASS	GLOBALNAJA NAWIGAZIONNAJA SPUTNIKOWAJA SISTEMA
GPS	Global Positioning System
HDD	Hard Disk Drive
HMI	Human-Machine-Interface
Hz	Hertz
ID	Identität
IT	Informationstechnologie
KDP	Knowledge Discovery Process
km	Kilometer
KW	Kürzeste Wege
LDW	Lane Departure Warning
LE	Letztmöglicher Eingriff
MiD2008	Mobilität in Deutschland 2008

MMS	Mensch-Maschine-Schnittstelle
MOST	Media Oriented Systems Transport
NDS	Naturalistic Driving Study
OEM	Original Equipment Manufacturers
PDA	Personal Digital Assistance
PG	Personalisiert - Gewichte
PK	Personalisiert - Kombiniert
PND	Portable Navigation Devices
PZ	Personalisiert - Zielfunktion
RU	Ruhig/gelassener Typ
SP	Sportlich/aggressiver Typ
SSQ	Sensation Seeking Questionnaire
ST	Standard
TMC	Traffic Message Channel
TV	Television
VICS	Vehicle Information and Communication System

Variablenverzeichnis

Variable	Bedeutung
ANT_f	Menge aller Anteilsabweichungen eines Fahrers
$ANT_{r,f}$	Anteilsabweichung einer Route eines Fahrers von einer FR-Route
$ANT_{AB_{r,f}}$	Anteil Autobahnkanten an einer Route eines Fahrers
$ANT_{IN_{r,f}}$	Anteil Innerortskanten an einer Route eines Fahrers
$ANT_{UB_{r,f}}$	Anteil Überlandkanten an einer Route eines Fahrers
ANT_{Uk_f}	Menge der Anteile an übereinstimmenden Kanten mit einer FR-Route eines Fahrers
$ANT_{Uk_{r,f}}$	Anteil übereinstimmender Kanten einer Route eines Fahrers mit einer FR-Route
$d_r(calc)$	Berechnete Routendauer
$d_{r,f}(obs)$	Gemessene Routendauer für einen Fahrer
$d_r(obs_th)$	Theoretische Dauer einer Route
$d_{r,f}(obs_th)$	Theoretische Dauer einer Route eines Fahrers
E	Anzahl Kanten für Laufzeitabschätzung
f	Fahrer
g	Geschwindigkeitsklasse
h	Stunden
G	Menge der Mittelwerte der Kantengewichte
G_{Sk}	Kantenbasierte Abweichung der Straßenklassenanteile von der FR-Route
G_{Sk_f}	Kantenbasierte Abweichung der Straßenklassenanteile von der FR-Route eines Fahrers
$G_{Sk}(vgl)$	Kantenbasierte Abweichung der Anteile an Straßencharakteristika von der PK-Route
$G_{Sk_f}(vgl)$	Kennwert der kantenbasierten Abweichung der Anteile an Straßencharakteristika von der PK-Route eines Fahrers

G_{Sl}	Längenbasierte Abweichung der Straßenklassenanteile von der FR-Route
G_{Sl_f}	Längenbasierte Abweichung der Straßenklassenanteile von der FR-Route eines Fahrers
$G_{Sl}(vgl)$	Längenbasierte Abweichung der Anteile an Straßencharakteristika von der PK-Route
$G_{Sl_f}(vgl)$	Kennwert der längenbasierten Abweichung der Anteile an Straßencharakteristika von der PK-Route eines Fahrers
G_{Uk}	Kantenbasierte Abweichung der Übereinstimmung von der PK-Route
G_{Uk_f}	Kennwert der kantenbasierten Abweichung der Übereinstimmung eines Fahrers von der PK-Route
$G_{Uk}(vgl)$	Kantenbasierter Vergleichswert der Übereinstimmung mit der PK-Route
$G_{Uk_f}(vgl)$	Kennwert des kantenbasierten Vergleichs der Übereinstimmung eines Fahrers mit der PK-Route
$g_Sk_{r,f}(vgl)$	Kantenbasierter Vergleichswert der Straßenklassenanteile einer Route eines Fahrers
G_Sk_f	Menge der kantenbasierten Vergleichswerte der Straßenklassenanteile eines Fahrers
$g_Uk_{r,f}(vgl)$	Kantenbasierter Übereinstimmungswert der Straßenklassenanteile einer Route eines Fahrers
G_Uk_f	Menge der kantenbasierten Übereinstimmungswerte der Straßenklassenanteile eines Fahrers
$g_Zd_{r,f}$	Einordnung der Routendauer einer Route eines Fahrers zur FR-Route
G_Zd_f	Menge der Einordnungen der Routendauer eines Fahrers zur FR-Route
$g_Zd_{r,f}(vgl)$	Vergleichswert der Routendauer einer Route eines Fahrers zur PK-Route
$G_Zd_f(vgl)$	Menge der Vergleichswerte der Routendauer eines Fahrers PK-Route
$g_Zl_{r,f}$	Einordnung der Routenlänge einer Route eines Fahrers zur FR-Route
G_Zl_f	Menge der Einordnungen der Routenlänge eines Fahrers zur FR-Route

$g_Zl_{r,f}(vgl)$	Vergleichswert der Routenlänge einer Route eines Fahrers zur PK-Route
$G_Zl_f(vgl)$	Menge der Vergleichswerte der Routenlänge eines Fahrers zur PK-Route
$G_Z(vgl)$	Abweichung der aggregierten Zielkriterien von der PK-Route
$G_{Z_f}(vgl)$	Kennwert der Abweichung der aggregierten Zielkriterien eines Fahrers von der PK-Route
G_{Zd}	Abweichung der Routendauer von der FR-Route
G_{Zd_f}	Kennwert der Abweichung der Routendauer eines Fahrers von der FR-Route
$G_{Zd}(vgl)$	Abweichung der Routendauer von der PK-Route
$G_{Zd_f}(vgl)$	Kennwert der Abweichung der Routendauer eines Fahrers von der PK-Route
G_{Zl}	Abweichung der Routenlänge von der FR-Route
G_{Zl_f}	Kennwert der Abweichung der Routenlänge eines Fahrers von der FR-Route
$G_{Zl}(vgl)$	Abweichung der Routenlänge von der PK-Route
$G_{Zl_f}(vgl)$	Kennwert der Abweichung der Routenlänge eines Fahrers von der PK-Route
$I_{Größe}$	Indikator der Größe
$I_{Größe_f}$	Indikatorwert der Größe eines Fahrers
$I_{GrößeMM_f}$	Min-Max-Indikatorwert der Größe eines Fahrers
$I_{Struktur}$	Indikator der Struktur
$I_{Struktur_f}$	Indikatorwert der Struktur eines Fahrers
$I_{StrukturMM_f}$	Min-Max-Indikatorwert der Struktur eines Fahrers
$inKern$	Menge der Knotenanzahlen des inneren Kerns
$I_Größe$	Menge der Indikatorwerte der Größe
$I_Struktur$	Menge der Indikatorwerte der Struktur
k	Kante
K_f	Kennwert des Min-Max-Vergleichs für einen Fahrer
K_g	Kennwert der Kantengewichte
K_{g_f}	Kennwert der Kantengewichte eines Fahrers
K_{inKern}	Kennwert des inneren Kerns
K_{inKern_f}	Kennwert des inneren Kerns eines Fahrers
K_{Kern}	Kennwert des Kerns
K_{Kern_f}	Kennwert des Kerns eines Fahrers
K_p	Kennwert der Knotenanzahl

K_{p_f}	Kennwert der Knotenanzahl eines Fahrers
K_q	Kennwert der Kantenanzahl
K_{q_f}	Kennwert der Kantenanzahl eines Fahrers
K_t	Kennwert der Kreisanzahl
K_{t_f}	Kennwert der Kreisanzahl eines Fahrers
$Kanten_{r,f}$	Menge der Kanten einer Route eines Fahrers
$Kanten_{AB_{r,f}}$	Menge der Autobahnkanten einer Route eines Fahrers
$Kanten_{IN_{r,f}}$	Menge der Innerortskanten einer Route eines Fahrers
$Kanten_{UB_{r,f}}$	Menge der Überlandkanten einer Route eines Fahrers
$Kanten_{Uk_{r,f}}$	Menge der übereinstimmende Kanten einer Route eines Fahrers mit einer FR-Route
$Kern$	Menge der Knotenanzahlen des Kerns
L	Maximales Gewicht einer Kante für Laufzeitabschätzung
$l_r(calc)$	Berechnete Routenlänge
$l_r(obs.th)$	Theoretische Länge einer Route
$l_{r,f}(obs.th)$	Theoretische Länge einer Route eines Fahrers
$l_{r,f}(obs)$	Gemessene Routenlänge
M_a	Maß der Abdeckung
M_{a_f}	Kennwert des Maßes der Abdeckung eines Fahrers
M_{aMM}	Maß der Abdeckung nach Min-Max-Normierung
M_{aMM_f}	Kennwert des Maßes der Abdeckung eines Fahrers nach Min-Max-Normierung
M_d	Maß der Dauer
M_{d_f}	Kennwert des Maßes der Dauer eines Fahrers
M_{dMM}	Maß der Dauer nach Min-Max-Normierung
M_{dMM_f}	Kennwert des Maßes der Dauer eines Fahrers nach Min-Max-Normierung
M_l	Maß der Länge
M_{l_f}	Kennwert des Maßes der Länge eines Fahrers
M_{lMM}	Maß der Länge nach Min-Max-Normierung
M_{lMM_f}	Kennwert des Maßes der Länge eines Fahrers nach Min-Max-Normierung
M_r	Maß der Dimension des Raumes
M_{r_f}	Kennwert des Maßes der Dimension des Raumes eines Fahrers
$M_{r_f}(Norm)$	Anteil der Dimension des Raumes an zweidimensionaler Zielfunktion für einen Fahrer

M_v	Maß der Geschwindigkeit
M_{v_f}	Kennwert des Maßes der Geschwindigkeit eines Fahrers
M_{vMM}	Maß der Geschwindigkeit nach Min-Max-Normierung
M_{vMM_f}	Kennwert des Maßes der Geschwindigkeit eines Fahrers nach Min-Max-Normierung
M_z	Maß der Dimension der Zeit
M_{z_f}	Kennwert des Maßes der Dimension der Zeit eines Fahrers
$M_{z_f}(Norm)$	Anteil der Dimension der Zeit an zweidimensionaler Zielfunktion für einen Fahrer
N	Anzahl
O	Landau-Symbol
p	Anzahl Knoten
P	Menge der Knotenanzahlen
p_f	Anzahl Knoten eines Fahrers
$p(inKern)$	Anzahl Knoten des inneren Kerns
$p_f(inKern)$	Anzahl Knoten des inneren Kerns eines Fahrers
$p(Kern)$	Anzahl Knoten des Kerns
$p_f(Kern)$	Anzahl Knoten des Kerns eines Fahrers
q	Anzahl Kanten
Q	Menge der Kantenanzahlen
q_f	Anzahl Kanten eines Fahrers
r	Route
$rd_{r,f}$	Abweichung der Routendauer einer Route eines Fahrers von einer Referenzroute
RD_f	Menge der Abweichungen der Routendauer für einen Fahrer
R_{inKern}	Relativwert des inneren Kerns
R_{inKern_f}	Relativwert des inneren Kerns eines Fahrers
R_{Kern}	Relativwert des Kerns
R_{Kern_f}	Relativwert des Kerns eines Fahrers
$RinK$	Menge der Relativwerte des inneren Kerns
RK	Menge der Relativwerte des Kerns
$rl_{r,f}$	Abweichung der Routenlänge einer Route eines Fahrers von einer Referenzroute
RL_f	Menge der Abweichungen der Routenlänge für einen Fahrer

$rv_{g,f}$	Abweichung der Geschwindigkeit einer Klasse für einen Fahrer
RV_f	Menge der Abweichungen der Geschwindigkeit eines Fahrers
t	Anzahl Kreise
T	Menge der Kreisanzahlen
t_f	Anzahl Kreise eines Fahrers
V	Anzahl Knoten für Laufzeitabschätzung
$v_g(map)$	Geschwindigkeitsklasse der digitalen Karte
$v_{g,f}(obs)$	Kantengewicht der Geschwindigkeit der Klasse eines Fahrers
$V_{g,f}$	Menge der gemessenen Geschwindigkeiten eines Fahrers
$v_{k,f}(obs)$	Gemessene Geschwindigkeit auf einer Kante für einen Fahrer

1 Einleitung

Navigationssysteme unterstützen den Menschen in seiner Mobilität, die als einer der Megatrends der Gegenwart und Zukunft gilt [57]. Für die individuelle Mobilität eines Fahrers bietet ein Navigationssystem eine umfassende *Navigationslösung* zur *Routenberechnung* an, die verschiedene Komponenten umfasst [67]. Somit ist es für den Fahrer möglich, jede Zielposition, auch in einem für ihn unbekannten Gebiet, zu erreichen.

Die Bewertung der berechneten Route durch den Fahrer als Nutzer des Navigationssystems basiert daher nicht darauf, ob die Zielposition erreicht wird, sondern welche *Güte* die Route besitzt. Die Güte kann z.B. durch die *Routendauer* oder die *Routenlänge* bestimmt werden [80]. Standardmäßig erfolgt die Routenberechnung statisch und ohne Einbeziehung des Fahrers und seiner *navigatorischen Präferenzen* [74]. Allerdings unterscheiden sich die Fahrer voneinander und haben mit ihrem Verhalten einen bedeutenden Einfluss auf die Güte der Route.

Das Ziel dieser Arbeit ist es, eine *Personalisierung* des Navigationssystems durchzuführen und somit den Fahrer und seine navigatorischen Präferenzen in die Routenberechnung einzubeziehen. Bisherige Ansätze hierzu werden in dieser Arbeit durch eine *Modellierung* des Fahrers erweitert, die verschiedene Aspekte seines *Routenauswahlverhaltens* umfasst.

Für die Modellierung ist es notwendig, Daten über den Fahrer und sein Routenauswahlverhalten zu erheben, diese zu analysieren und Wissen über den Fahrer zu generieren. Zusätzlich zu Methoden der *Wirtschaftsinformatik* werden Werkzeuge und Verfahren der *Statistik* und der *Graphentheorie* zur Wissensgenerierung genutzt. Ebenfalls bauen die Analysen auf den Erkenntnissen der *Verkehrspsychologie* auf. Hierfür wurden in zahlreichen Studien Unterschiede zwischen den Fahrern und ihrem Verhalten untersucht [16, 26].

Somit ist das in dieser Arbeit entwickelte Verfahren interdisziplinär angelegt. Es leistet damit einen Beitrag für eine umfassende Modellierung des Fahrers und wird in dieser Arbeit für die Personalisierung genutzt. Auf diese Weise wird der individuellen Mobilität des Fahrers durch eine individuelle Navigationslösung entsprochen

mit dem Ziel, die Güte aber auch die Akzeptanz heutzutage verfügbarer Navigationssysteme zu erhöhen.

Dafür ist es notwendig, dass die Personalisierung für das in heutzutage verfügbaren Navigationssystemen genutzte Verfahren anwendbar ist. Das Verfahren besteht seit der Entwicklung von Navigationssystemen aus den gleichen Komponenten und ist wie folgt aufgebaut: Grundlage ist die Positionierung des Fahrzeugs auf dem in einer *digitalen Karte* gespeicherten Straßennetz (*Ortung*) [92]. Das gespeicherte Straßennetz wird durch einen *Graphen* [90] bestehend aus *Knoten* und *Kanten* repräsentiert. Den Kanten sind Attribute wie Länge oder angenommene Geschwindigkeit zugeordnet [86].

Die Routenberechnung bzw. das *Routing* ist der Kern eines jeden Navigationssystems und bestimmt eine Route zu einer vom Fahrer angegebenen Zielposition in Abhängigkeit von einem *Zielkriterium*. Anschließend wird der Fahrer entlang der Route zur Zielposition geführt (*Routenführung*).

Für die Berechnung der Route wird ein *Kürzeste-Wege-Verfahren (KW-Verfahren)* [22, 55, 84] genutzt, das die Attribute der Kanten als Gewichte verwendet und eine minimale Route für eine *Zielfunktion* berechnet. Die Minimierung der Routendauer oder der Routenlänge sind die am Häufigsten für die Zielfunktion genutzten Zielkriterien [80]. Dabei ist die Routendauer der *Dimension der Zeit* und die Routenlänge der *Dimension des Raumes* zugeordnet. Jede Route kann bezüglich dieser beiden Dimensionen charakterisiert werden.

Die Kantengewichte ergeben sich aus den entsprechenden Attributen. Standardmäßig haben die Kanten für alle Fahrer einheitliche Gewichte [86]. Allerdings bevorzugen Fahrer unterschiedliche Geschwindigkeiten, so dass sich diese zum Teil stark von denen in der digitalen Karte gespeicherten Kantengewichten der Geschwindigkeit unterscheiden.

In heutzutage verfügbaren Navigationssystemen ist es daher möglich, dass die Routenberechnung automatisch oder manuell angepasste Geschwindigkeiten nutzt. Allerdings differenzieren diese Ansätze nur wenige Klassen und beziehen die Dimension des Raumes nicht mit ein. So könnten zwei Kanten auf die gleiche Weise angepasst werden, sich die eine allerdings im städtischen und die andere im ländlichen Gebiet befinden.

Zudem hat die Wahl der Geschwindigkeit auch Einfluss auf die vom Fahrer gewählten Zielkriterien. Einige Fahrer bevorzugen schnellere, andere Fahrer kürzere Routen, wobei Fahrer keine reinen Zielkriterien nutzen. Veranschaulichen lässt sich dies am Beispiel einer Route, die wenige Sekunden schneller, dafür aber mehrere Kilo-

meter länger als eine Alternativroute ist. Eine Routenberechnung mit dem alleinigen Zielkriterium Routendauer würde die erste, ein menschlicher Fahrer allerdings die Alternativroute auswählen. Bei anderen, sich nicht so eindeutig unterscheidbaren Routen, kann es dabei zu Unterschieden zwischen den Fahrern kommen.

Für die Routenberechnung können Unterschiede zwischen den Fahrern durch eine personalisierte Gewichtung der beiden Kriterien abgebildet werden. Dies ist bisher nur auf manuelle Weise für die Zielfunktion möglich. Hierfür wird ein Verhältnis von Routendauer und Routenlänge und damit der Dimensionen der Zeit und des Raumes durch den Fahrer angegeben. Als Folge werden kombinierte Kantengewichte für die Berechnung genutzt. Allerdings ist dieser mit Aufwand verbundene manuelle Ansatz fehleranfällig und steht in keinem direkten Zusammenhang zu den Kantengewichten der Geschwindigkeit.

Somit fehlt in den bisherigen Ansätzen eine Modellierung des Zusammenhangs zwischen den Dimensionen der Zeit und des Raumes und wie die Dimension der Zeit vom Fahrer durch seine gewählten Geschwindigkeiten gewichtet wird. Dieses bildet allerdings die Grundlage bei der Auswahl einer Route durch den Fahrer. Seine navigatorischen Präferenzen werden nur bedingt und nicht gesichert in die Routenberechnung einbezogen, was zu einer vom Fahrer unerwünschten Route führen bzw. die Zielkriterien des Fahrers in geringerem Maße erfüllen kann.

Daher wird in dieser Arbeit eine automatisierte und umfassende Einbeziehung des Fahrers für eine Personalisierung der Routingkomponente von Navigationssystemen durchgeführt. Das Ziel ist die Berechnung von Routen, die in größerem Maße den navigatorischen Präferenzen des Fahrers entsprechen und eine bessere Zielerreichung wie auch eine höhere Akzeptanz zur Folge haben.

Das Vorgehen ist an einen *Knowledge Discovery Process (KDP)* [27] angelehnt und wurde auf Basis eines umfangreichen Realfahrdatensatzes entwickelt. Dieser Datensatz, erhoben im Projekt *euroFOT* [47], umfasst ca. 100 Fahrer, denen für drei Monate ein Fahrzeug zur Verfügung gestellt wurde. Bei der Nutzung des Fahrzeugs war ihnen die Wahl der Zielposition sowie der Route vollkommen freigestellt, so dass ihr Routenauswahlverhalten nicht beeinflusst wurde und für die Analyse genutzt werden kann. Für die Daten werden zunächst mehrere Aufbereitungsschritte durchgeführt. Nach der Datenauswahl werden die Daten bereinigt. Der Schwerpunkt bei der Datenbereinigung liegt auf den Positionen und ihrer Zuordnung zu einer Kante des Straßennetzes. Abschließend werden die Daten für die Analyse drei verschiedener Perspektiven (*Kanten-, Routen- und Netzwerkperspektive*) reduziert.

Für die Kantenperspektive werden die gemessenen Geschwindigkeiten der Fahrer auf den verschiedenen Geschwindigkeitsklassen untersucht. Damit ist diese Perspektive der Dimension der Zeit zugeordnet. Die Ergebnisse der Analysen werden sowohl für die Modellierung wie auch für die Anpassung der Kantengewichte der Geschwindigkeit bei der Personalisierung genutzt.

Die vom Fahrer selbstständig ausgewählten Routen (*Fahrerrealrouten*) werden für die Routenperspektive analysiert. Dafür werden mit Hilfe einer neutralen *externen Routingreferenz* für die gemessenen Start-Zielkombinationen *Referenzrouten* berechnet. Die für diese Routen ermittelten Routendauern und -längen werden mit den gemessenen verglichen. Die berechneten Routen werden somit als Benchmark genutzt und ermöglichen eine Einordnung der Fahrer bezüglich der Dimension der Zeit (Routendauer) wie auch der Dimension des Raumes (Routenlänge). Beide Einordnungen gehen in das Modell ein.

Die Netzwerkperspektive betrifft das vom Fahrer in den drei Monaten gezeigte Mobilitätsverhalten. Das auf diese Weise aufgespannte Netz wird durch einen Graphen repräsentiert und mit graphentheoretischen Analysewerkzeugen untersucht. Das Ergebnis ist eine Einordnung des Fahrers, die der Dimension des Raumes zugeordnet werden kann. Insbesondere diese Einordnung ist wichtig für die Erweiterung der Modellierung gegenüber bisherigen Ansätzen.

Das Ergebnis der Analysen der drei Perspektiven sind jeweils zwei Modelle für beide Dimensionen. Deren Zusammenhang wird zunächst statistisch untersucht. Anschließend werden die Modelle aggregiert, um so ein fundiertes, zweidimensionales Modell, das sowohl die Dimension der Zeit als auch die Dimension des Raumes umfasst, zu erhalten. Es bildet die Grundlage für die Personalisierung der Routenberechnung.

Für die Berechnung der personalisierten Routen wird eine *interne Routingreferenz* genutzt. Die dazugehörige digitale Karte basiert auf den erhobenen Kanten des Realfahrdatensatzes. Damit ist die Datengrundlage leichter veränderbar, vor allem aber auch vergleichbar zu den gemessenen Routen der Fahrer. Es werden drei verschiedene Ansätze der Personalisierung durchgeführt: *Personalisierung der Kantengewichte*, *Personalisierung der Zielfunktion* sowie eine *kombinierte Personalisierung*.

Für die Personalisierung der Gewichte werden ausschließlich die Kantengewichte der Geschwindigkeit angepasst und auf dieser Basis eine schnellste Route berechnet. Es wird bei diesem Ansatz nur die Dimension der Zeit berücksichtigt. Bei der Personalisierung der Zielfunktion werden für jeden Fahrer Anteile für die Dimensi-

on der Zeit und der Dimension des Raumes für eine zweidimensionale Zielfunktion ermittelt. Die Routenberechnung wird allerdings auf Basis der Standardkantengewichte durchgeführt. Somit werden zwar beide Dimensionen mit einbezogen, doch werden die Geschwindigkeiten und damit die Gewichtung der Dimension der Zeit für jeden Fahrer nicht berücksichtigt. Für die kombinierte Personalisierung werden beide vorherigen Ansätze kombiniert: Für eine personalisierte zweidimensionale Zielfunktion wird eine Route auf Basis der angepassten Kantengewichte berechnet. Somit werden beide Dimensionen differenziert mit einbezogen.

Die verschiedenen Ansätze werden untereinander sowie mit einer *Standardroute* verglichen. Dies ist eine schnellste Route, die auf Basis standardisierter Kantengewichte der erstellten digitalen Karte mit Hilfe der internen Routingreferenz berechnet wird. Da es Ziel der Personalisierung ist, Routen zu berechnen, die eine höhere Güte aufweisen als die Standardroute, ist es notwendig diese zu messen. Dafür werden drei verschiedene *Gütemaße* genutzt: *Zielkriterien*, *Streckencharakteristika* und *Übereinstimmung*.

Das erste Maß der Zielkriterien analysiert und vergleicht die Routendauer und die Routenlänge der berechneten Routen. Dazu wird eine Detailanalyse für die Einflüsse der verschiedenen Perspektiven durchgeführt. Für das Maß der Streckencharakteristika werden die Straßenklassenanteile der berechneten Routen mit denen der Fahrerrealrouten verglichen und die Veränderung analysiert. Das dritte Maß der Übereinstimmung untersucht, inwieweit die berechneten Routen mit der gemessenen Route übereinstimmen.

Mit Hilfe der Gütemaße lässt sich der Nutzen der modellgestützten Personalisierung quantifizieren. Insbesondere werden der Einfluss der Dimension des Raumes und die Rolle der Maße der verschiedenen Perspektiven untersucht mit dem Ziel, die Routenberechnung von heutzutage verfügbaren Navigationssystemen zu verbessern.

Die Arbeit ist folgendermaßen aufgebaut. In Kapitel 2 werden die Grundlagen der Navigation dargestellt. Diese beinhalten die historische und technische Entwicklung von Navigationssystemen, ihre Funktionsweise und Komponenten sowie eine Analyse des Marktes für Navigationslösungen. Dazu werden sowohl umgesetzte Verfahren zur Personalisierung als auch der Stand der Forschung dargestellt. So werden das Potenzial aber auch bestehende Restriktionen für Ansätze zur Verbesserung der Routingkomponente von Navigationssystemen aufgezeigt.

Aufbauend auf diesen Erkenntnissen werden in Kapitel 3 das Konzept und der Ansatz des in dieser Arbeit entwickelten Verfahrens für die Modellierung und die

Personalisierung erläutert. Der Aufbau und die Schritte eines KDP als Grundlage des Verfahrens werden ebenfalls beschrieben. Sowohl das Konzept wie auch das Verfahren werden empirisch mit Hilfe des Realfahrdatensatzes belegt. Der für das Verfahren genutzte Datensatz aus dem Projekt euroFOT wird in Kapitel 4 dargestellt.

Das entwickelte Verfahren für die Modellierung und die Personalisierung ist Thema von Kapitel 5. Die Modellierung umfasst das Vorgehen bei den einzelnen Perspektiven sowie die Aggregation. Die verschiedenen Personalisierungsansätze werden ebenfalls explizit vorgestellt sowie das Verfahren zur Bewertung erläutert. Die Datenaufbereitung des Realfahrdatensatzes wird in Kapitel 6 beschrieben.

Nach der Aufbereitung werden die empirischen Ergebnisse der Modellierung und der Personalisierung in Kapitel 7 dargestellt. Es umfasst die Ergebnisse der Perspektivenanalysen, die zweidimensionale Modellierung sowie die Personalisierung. Ein weiterer Schwerpunkt liegt auf der Bewertung der Personalisierung. In Kapitel 8 wird ein Fazit der Arbeit gezogen und ein Ausblick gegeben.

2 Grundlagen der Navigation

2.1 Navigationssysteme

2.1.1 Entwicklung der Navigation

Die Entwicklung des Automobils stellte für die individuelle Mobilität des Menschen einen enormen Fortschritt dar und ermöglichte einer zunehmenden Anzahl an Menschen ein umfangreiches Mobilitätsverhalten [57]. Dabei war und ist es notwendig, den Fahrer bei der dazugehörigen Navigationsaufgabe zu unterstützen, wenn er sich in einem ihm unbekannten Gebiet bewegt.

Für diese Aufgabe bildeten entsprechend detaillierte Straßenkarten die Grundlage. Die Bestimmung der eigenen Position (Ortung), die Auswahl einer geeigneten Route vom Start- zu einer Zielposition (Routenberechnung) sowie das Folgen der ausgewählten Route (Routenführung) waren allerdings auch weiterhin Aufgabe des Fahrers. Durch die Entwicklung von Navigationssystemen zu Beginn der 1980er Jahre wurden dem Fahrer auch diese Aufgaben abgenommen und stellten eine durchgehende Navigationslösung zur Unterstützung der Mobilität zur Verfügung [67].

Für eine der größten technischen Herausforderungen, der korrekten Positionierung des Fahrzeugs, waren verschiedene Sensoren notwendig, um die Position des sich bewegenden Fahrzeugs, ausgehend von einer bekannten Position, fortzuschreiben. Sensoren, wie Gyrometer für die Drehbewegung, messen die Beschleunigungen und Verlagerungen des Fahrzeugs und bestimmen auf einer gespeicherten digitalen Karte die Position [67]. Die Karte besaß keinen hohen Detailgrad, da Anwendungen im zivilen Bereich aufgrund der Komplexität und der damit verbundenen Kosten nicht weit verbreitet waren. Ein Navigationssystem musste immer im Fahrzeug integriert (*fest eingebautes Navigationssystem*) sein, um Zugriff auf die Daten der benötigten Sensoren zu haben.[67]

Eine weitreichende Änderung erfuhr die Navigation durch die Freigabe des amerikanischen *Global Positioning Systems* (GPS) für private Nutzung in den neun-

ziger Jahren [66]. Somit war es möglich, aufgrund von Satellitensignalen und darauf aufbauender Berechnungen die Position metergenau zu bestimmen. Da sich Fahrzeuge im Allgemeinen auf dem Straßennetz befinden und sich darauf bewegen, war es somit möglich, eine sehr genaue Ortung durchzuführen ohne aufwendige Sensorik zu benötigen. Durch die Einsparpotenziale kosteten Navigationssysteme nur noch wenige Tausend D-Mark und wurden für den privaten Nutzer erschwinglich. Karten wurden verbessert und benutzerfreundliche Bedienungs- und Routenführungskonzepte entwickelt [67]. Durch höhere Stückzahlen konnte günstiger produziert werden und weitere Nutzer gewonnen werden, was das Wachstum verstärkte [70]. Die Absatzzahlen der Navigationssysteme in Europa stiegen von 1997 bis 2002 von unter 100 000 auf über 1.2 Millionen Stück stark an, siehe Abbildung 2.1.

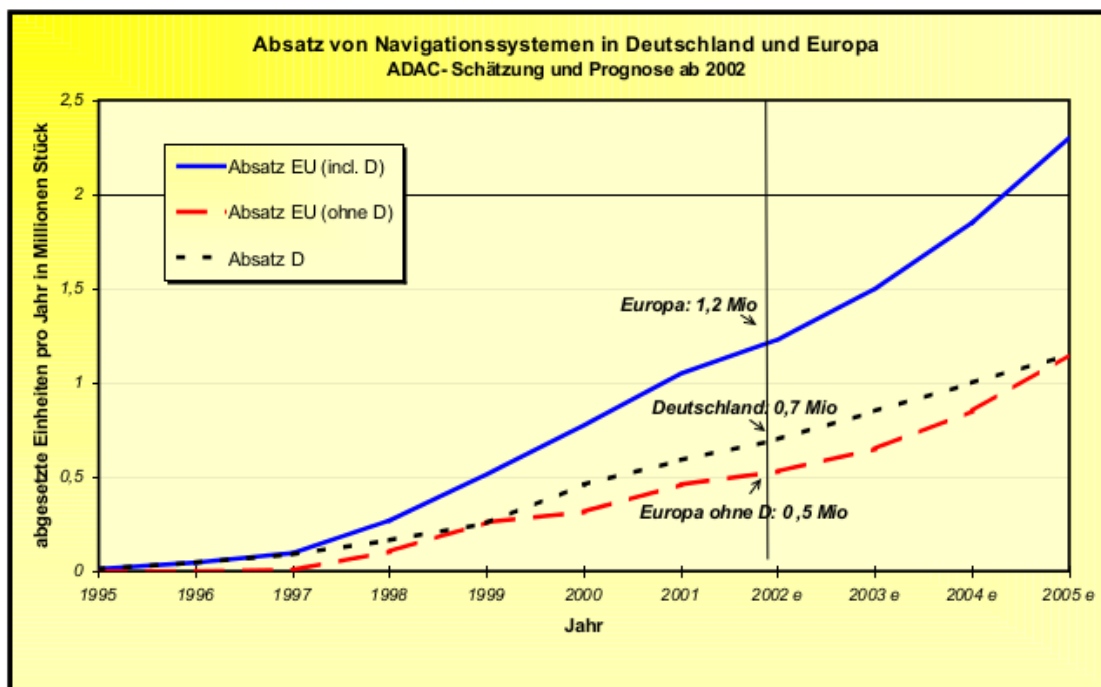


Abbildung 2.1: Absatzzahlen von festeingebauten Navigationssystemen in Deutschland und Europa, Abbildung aus [70]

Anfangs wurden viele Fahrzeuge nachgerüstet, so dass es im Jahre 1998 etwa gleich viele Erstausrüstungen der *Original Equipment Manufacturers* (OEM) mit Navigationssystemen ab Werk und Nachrüstungen im sogenannten *Aftermarket* (AM) gab. Dieses Verhältnis änderte sich auf 70% zu 30% zugunsten der Erstausrüstungen. In der Abbildung 2.1 aus dem Jahre 2002 wurde davon ausgegangen, dass der Markt insgesamt weiter stark ansteigt und sich das Verhältnis weiter zugunsten der Erstausrüstung verschiebt [70].

Durch den technischen Fortschritt im Bereich der Speicher- und Prozessortechnik war es auch möglich, vom Fahrzeug unabhängige Navigationssysteme ausschließlich auf GPS basierend zu entwickeln. Diese mobilen Navigationssysteme bzw. *Portable Navigation Devices* (PND) stellten eine große Neuerung im Bereich der Navigation dar [67]. Sie werden vornehmlich an der Frontscheibe befestigt, um einen guten GPS-Empfang sowie dem Fahrer eine gute Sicht auf das Display für die Routenführung zu ermöglichen. Durch die deutliche Verringerung der benötigten Sensoren und Komponenten waren diese Systeme deutlich günstiger als festeingebaute Systeme für wenige Hundert Euro zu erwerben und vergrößerten den potentiellen Markt für Navigationssysteme enorm [100].

PNDs verbreiteten sich ab dem Jahr 2005 stark (siehe Abbildung 2.2) und ermöglichten so vielen Fahrern das Navigieren auf unbekannten Strecken ohne manuelles Auswählen einer Route. Als Folge der Verbreitung wurden auch die digitalen Karten verbessert und standardisiert. Dabei setzten die zwei Kartenhersteller *TeleAtlas* und *Navtech* den Standard und bilden, trotz Übernahme durch andere Firmen, bis heute mit ihren Karten die Grundlage für die meisten Navigationssysteme [20].

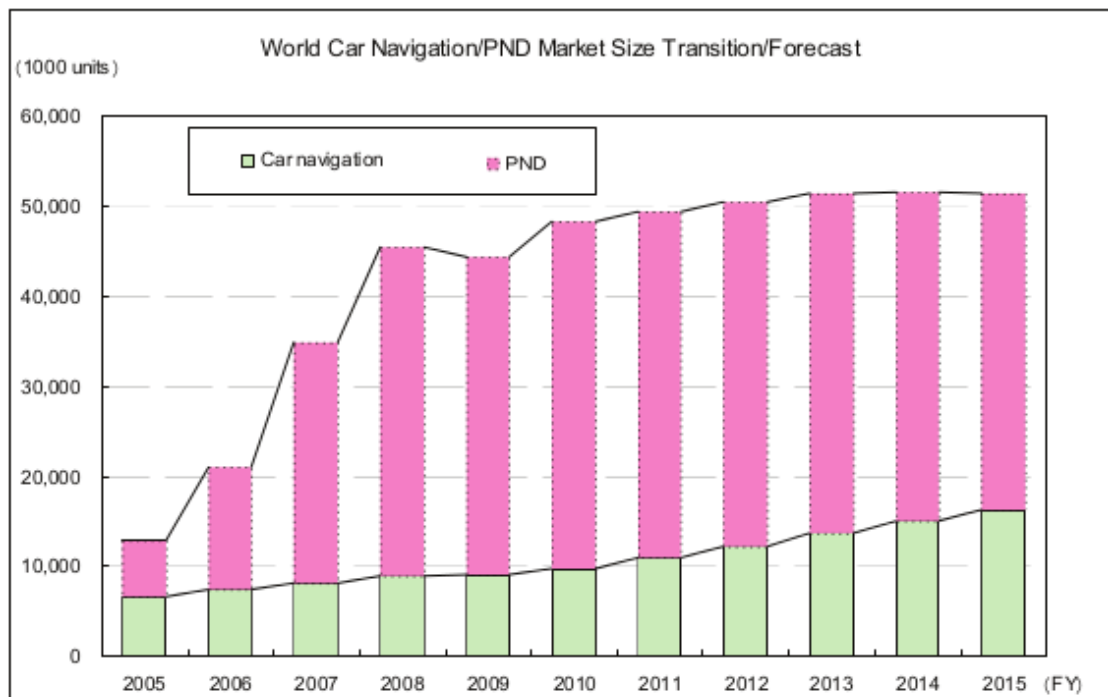


Abbildung 2.2: Absatzzahlen von Navigationssystemen weltweit bis 2008, Prognose der Absatzzahlen bis 2015, Abbildung aus [100]

Durch die deutlich geringeren Entwicklungszeiten profitierten auch hier mobile Navigationssysteme von Verbesserungen der digitalen Karte. Dementsprechend war

2008 der Marktanteil von verkauften mobilen Navigationssystemen auf gut 80% gestiegen [100].

Es wurden weitere Funktionen für die Navigation entwickelt sowie unterschiedliche Designs erstellt. Eine wichtige Komponente stellt das *Human-Machine-Interface* (HMI) dar, das die berechnete Route möglichst detailliert und anschaulich darstellt. Auch hier hatten PNDs Vorteile durch kürzere Entwicklungszeiten. Allerdings wurden durch den Druck der PNDs auch bei den festeingebauten Navigationssystemen weitere Entwicklungen getätigt, um einen Mehrwert zu generieren und den weiterhin höheren Preis zu begründen [67]. Eine zusammenfassende Darstellung der wichtigsten technischen Entwicklungen von 1980 bis 2010 ist in Abbildung 2.3 zu finden. Daher wurde im Jahr 2009 davon ausgegangen, dass PNDs weiter deutlich vorherrschend sind, der Absatz von festeingebauten Systemen sich aber steigert und allmählich Marktanteile zurückgewinnt, siehe Abbildung 2.2 [100].

1981	• Internal navigation system with gas-rate sensor
1985	• Commercial-type navigation system that used in digital electric map
1987	• DR type car navigation system with geomagnetic sensor and CD-ROM
1988	• Developed map-matching technology
1990	• GPS type car navigation system with CD-ROM
1991	• Route planning that is showed the way to destination
1992	• Route guidance with sound and voice at turn by turn
	• VICS system development began
1996	• VICS system service started
1997	• DVD-type in-car navigation system
	• All in-one audio & video in-car navigation
1999	• Hands free telephony and data communication using cellar phone system
2000	• GPS selective availability turned off
2001	• HDD type in-car navigation system
2002	• Communication module built-in type In-car navigation system
2003	• To support car probe system using cellular phone
2004	• Portable navigation device with flash memory
2005	• Linking portable music player type in-car navigation system
2006	• To support digital TV
2009	• To link Blu-Ray for automotive grade

Abbildung 2.3: *Entwicklungen im Bereich der Navigation, Abbildung nach [67]*

Weitere Entwicklungen, wie z.B. die Einbeziehung des aktuellen *Verkehrs* (siehe Abschnitt 2.3.2), erhöhten den Nutzen von Navigationssystemen [24, 60]. Dazu wurden Möglichkeiten evaluiert, wie Daten von Sensoren im Fahrzeug dem Navigationssystem zur Verfügung gestellt werden können, um diese durch zusätzliche Informationen weiter zu verbessern [67]. Trotz der Weiterentwicklungen und Zusatzkomponenten hat sich das Vorgehen beim Kern der Navigation, der Routenberechnung

nung nicht geändert. Auf Basis einer gespeicherten digitalen Karte, die durch einen Graphen repräsentiert wird, wird ein kürzester Weg berechnet, unabhängig davon, ob es sich um mobile oder festeingebaute Navigationssysteme handelt.

2.1.2 Markt für Navigationslösungen

Der Markt für PNDs wird von zwei Herstellern beherrscht. Dies sind der niederländische Hersteller *TomTom* und der amerikanische Hersteller *Garmin* [44]. War TomTom längere Zeit führend so ist seit 2007 Garmin Marktführer [4]. 2007 verkaufte Garmin 1.85 Millionen mobile Navigationssysteme und TomTom 1.81 Millionen [44]. 2008 war der Marktanteil von Garmin nochmals gestiegen und betrug 33.7 %, dicht gefolgt von TomTom mit 30% [76]. Daneben gibt es noch weitere Marktteilnehmer, wie z.B. *Navman* oder *Magellan*, die sich das restliche Drittel teilen.

Seitdem behauptet Garmin diese Position und konnte seinen Vorsprung systematisch ausbauen. So betrug der Jahresumsatz von Garmin im Jahre 2013 2.6 Milliarden Dollar (2 Milliarden Euro) [29], während TomTom 960 Millionen Euro umsetzte [30]. Der Umsatz wird mit Navigationssystemen für Fahrzeuge, aber auch neuen Produkten für andere Bereiche, wie Navigationssysteme für Freizeitaktivitäten (Laufen, Fahrradfahren, Wandern), generiert. Diese Systeme beinhalten weitere Funktionen, wie z.B. Pulsmessung oder Videoaufnahmefunktion, um dem Nutzer Daten und Bewertungen über die Strecke wie auch über sich selbst anzubieten [9].

Dazu ist der Markt für Navigationssysteme im Automobilbereich erneut im Umbruch, dem sich sowohl die großen wie auch die kleinen Anbieter stellen müssen [44, 77]. Im Bereich des Mobilfunks werden zunehmend *Smartphones* verkauft, siehe Abbildung 2.4. Von 2008 bis 2010 hatte sich der Umsatzanteil der Smartphones im Telekommunikationsbereich von 13% auf 58% mehr als vervierfacht [36].

Diese Entwicklung hat sich in den letzten Jahren fortgesetzt [39]. Smartphones sind nicht nur internetfähig sondern haben in der Regel auch einen GPS-Empfänger. Mit der mittlerweile erreichten Displaygröße, einer Halterung im Fahrzeug und einer entsprechenden (meist kostenlosen) App, z.B. navfree, gibt es nur noch wenig Unterschiede zu einem PND. Dazu hat der Mensch bzw. Fahrer sein Smartphone nahezu immer dabei, während das mobile Navigationssystem extra mitgeführt werden muss. Damit stellen die Smartphones eine große Bedrohung für den Markt der PNDs dar.

Umsatz Telekommunikation (Privatkäufe)

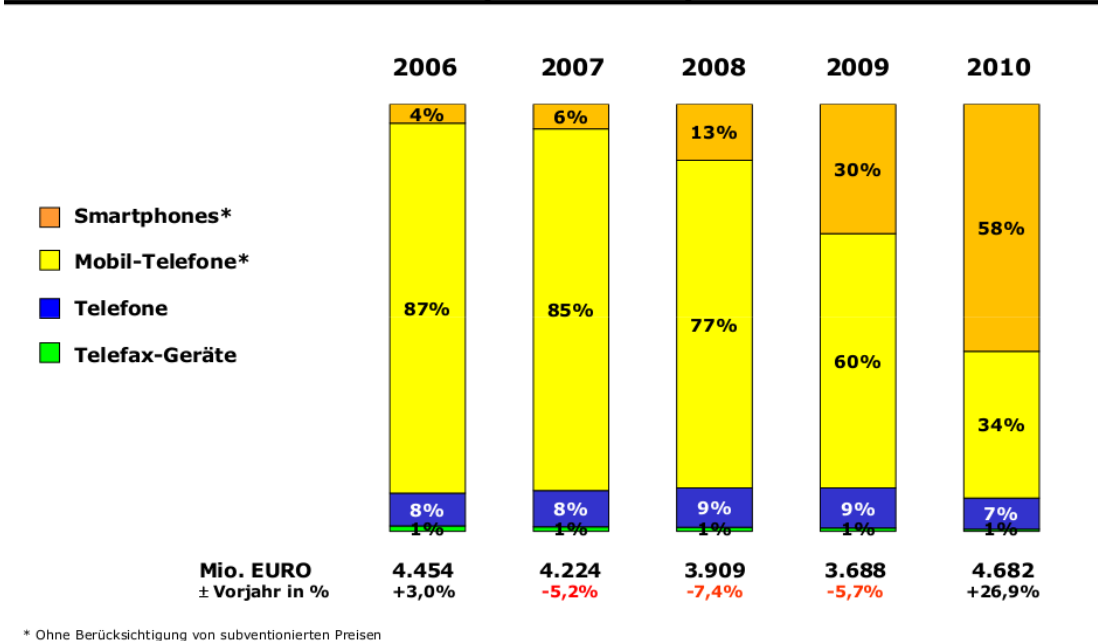


Abbildung 2.4: Absatzzahlen im Mobilfunkbereich, Abbildung aus [36]

Dies zeigt sich auch an den erreichten Umsätzen im Navigationsmarkt, siehe Abbildung 2.5. Die Hersteller von PNDs machten sich lange gegenseitig durch günstige Angebote Konkurrenz, die aber durch ein großes Marktwachstum ausgeglichen wurde.

Da der Markt nach den Jahren mit großem Wachstum gesättigt ist bzw. durch Smartphones bedroht wird, geht der Umsatz mit PNDs zurück. Betrug der Umsatz 2006 noch über 1.2 Milliarden Euro, so ist er bis 2010 auf unter 800 Millionen Euro gesunken. Daher geht die Entwicklung im mobilen Bereich zunehmend in die Richtung, dass das Gerät zur Navigation selber keinen Wert darstellt, sondern die damit verbundenen Dienste, wie Aktualität der digitalen Karte oder Erhebung des aktuellen Verkehrs, den Mehrwert darstellen.

Die Hersteller im AM stehen dabei vor den gleichen Problemen, zumal der Preis für solche Systeme mit über Tausend Euro deutlich höher ist. Der Bereich der OEMs kann sich dagegen besser von der Konkurrenz durch Smartphones absetzen und so seine Stellung halten. Diese Systeme sind ab Werk mit dem Fahrzeug integriert und können auf Daten der Fahrzeugsysteme zugreifen. Dazu kann der Preis durch Rabatte und Aktionen gesenkt werden.

Umsatz Car Audio/Navigation*

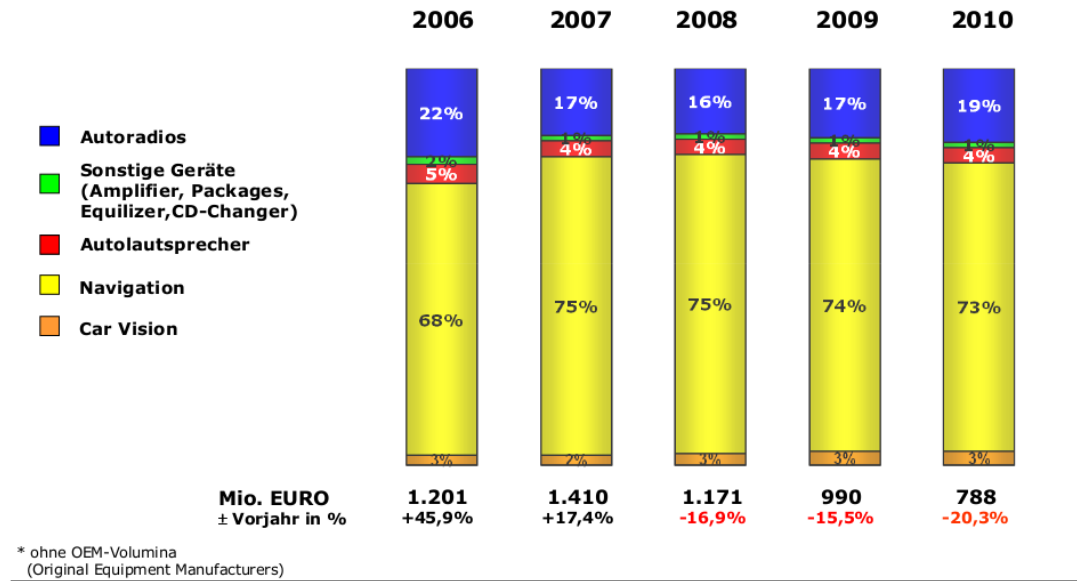


Abbildung 2.5: Umsatz im AM-Bereich, Abbildung aus [36]

Die Entwicklungen der Absatzzahlen sowie des Umsatzes für verschiedene Navigationssysteme sind in den Abbildungen 2.6 und 2.7 dargestellt. Der globale Navigationsmarkt ist von 2007 bis 2014 kontinuierlich von 50 Millionen auf 180 Millionen Stück gewachsen. Auch wenn PNDs weiterhin die knapp die Hälfte der abgesetzten Systeme ausmachen, so steigt der Anteil an Smartphones stetig an. Der Absatz festeingebauter Systeme ist ebenfalls gewachsen, allerdings nicht so stark wie der Absatz des Gesamtmarktes. Als Folge liegt der Marktanteil bei 10 bis 20%.

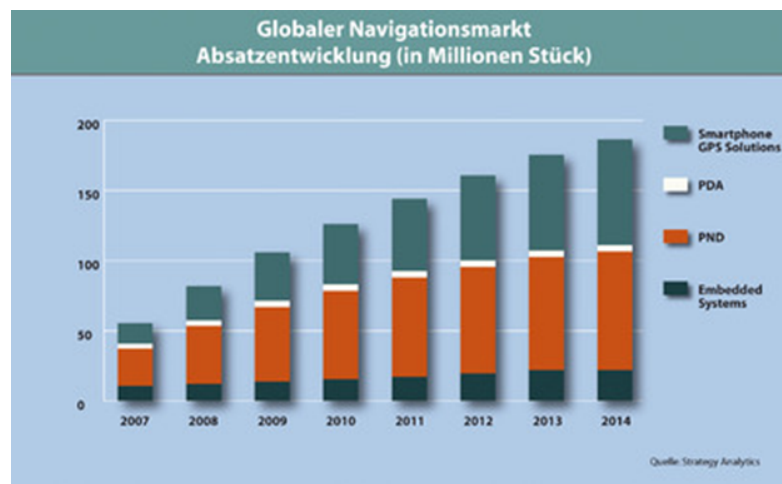


Abbildung 2.6: Absatzzahlen im globalen Markt für Navigation, Abbildung aus [94]

Der Absatz von mobilen tragbaren Computern (*Personal Digital Assistance (PDA)*) ist zwar ebenfalls aufgeführt, ihr Anteil am Navigationsmarkt ist allerdings verschwindend gering. Daher werden diese Geräte in der weiteren Betrachtung nicht berücksichtigt.

Bei der Umsatzentwicklung des Marktes ist eine zunehmende Sättigung bei 33 Milliarden US-Dollar sichtbar. Die steigenden Absatzzahlen sind nicht mit steigenden Umsätzen verbunden. Während der Umsatz der festeingebauten Systeme leicht wächst und der Anteil am Umsatz in etwa gleich bleibt, ging der Anteil der PNDs seit 2009 zunehmend zurück [94]. Dies ist durch den zunehmenden Preisverfall bei PNDs zu erklären. Zeitgleich stiegen der Umsatz für Smartphones sowie der Anteil am Umsatz des globalen Navigationsmarktes an. Es ist anzunehmen, dass diese Entwicklung sich fortsetzen und der Umsatz der PNDs weiter zurückgehen wird. Dazu zeigen die unterschiedlichen Anteile von festeingebauten Systemen an Absatz und Umsatz auch den deutlichen preislichen Unterschied zwischen diesen Systemen und PNDs bzw. Smartphones auf.

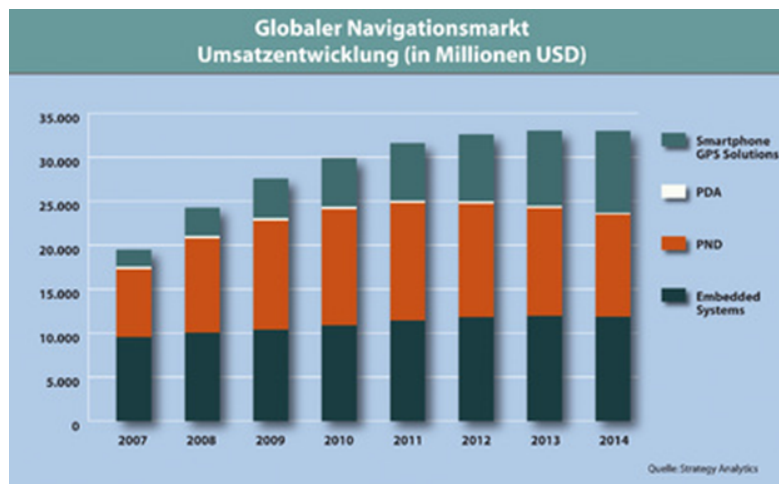


Abbildung 2.7: Umsatzzahlen im globalen Markt für Navigation, Abbildung aus [94]

Aufgrund der verringerten Umsätze bei PNDs versuchen die Hersteller ihre Produkte in Smartphones anzubieten bzw. zusätzliche Dienste zu integrieren, um damit einen Mehrwert zu generieren [9]. Damit werden alle Marktteilnehmer Anbieter von Navigationslösungen unterschiedlichster Art und Zusammensetzung. Gleichzeitig ist der generelle Bedarf an Navigationslösungen für den Menschen auch heutzutage hoch und bietet auch zukünftig ein sehr großes Potenzial.

2.2 Komponenten von Navigationssystemen

Aufgabe eines Navigationssystems ist es, eine Route von einem Start- zu einem Zielpunkt zu bestimmen und den Fahrer entlang dieser Route zum Zielpunkt zu führen. In Abbildung 2.8 ist der vereinfachte Aufbau eines Navigationssystems dargestellt. Anhand eines Satellitensignals wird die aktuelle Position bestimmt (Ortung). Diese geortete Position wird einer Position auf dem in der digitalen Karte gespeicherten Straßennetz zugeordnet (*Map Matching*) und der Routenberechnung bzw. dem Routing zur Verfügung gestellt. Die Routenberechnung bestimmt eine Route zu einem, vom Fahrer über die Bedienung eingegebenen, Zielpunkt. Der Startpunkt ist in der Regel die aktuelle Position. Mit Hilfe der Routenführung erreicht der Fahrer den Zielpunkt. Während sich das Fahrzeug bewegt, wird die jeweils aktuelle Position der Routenführung zur Verfügung gestellt. Die Routenführung sowie die Bedienung gehören zum HMI, das die Interaktion des Fahrers mit dem Navigationssystem ermöglicht.

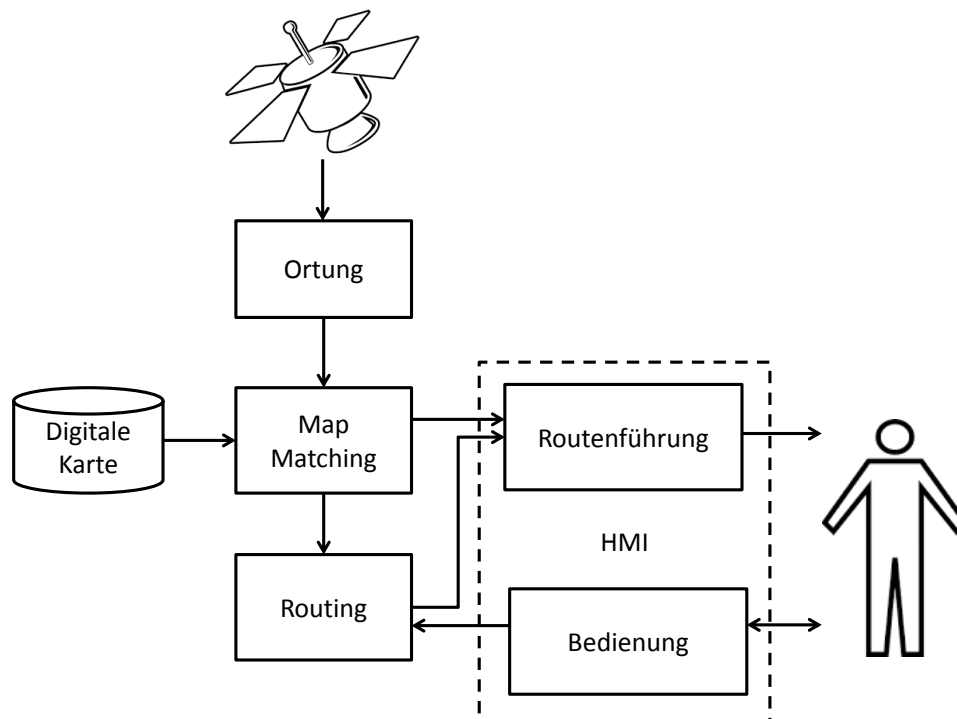


Abbildung 2.8: Vereinfachter Aufbau von Navigationssystemen

2.2.1 Digitale Karte

In jedem Navigationssystem im Fahrzeug ist eine digitale Karte gespeichert. Diese bildet die Grundlage für die Routenberechnung. Zentrale Elemente sind die Straßen des Straßennetzes, die als Kanten eines Graphen gespeichert werden [31]. Ein Graph besteht aus Knoten und die sie verbindenden Kanten. Eine Veranschaulichung ist in Abbildung 2.9 dargestellt. Die Kanten können gerichtet sein und verschiedene Attribute besitzen. Für eine detaillierte Erläuterung von Graphen, siehe [90, 93].

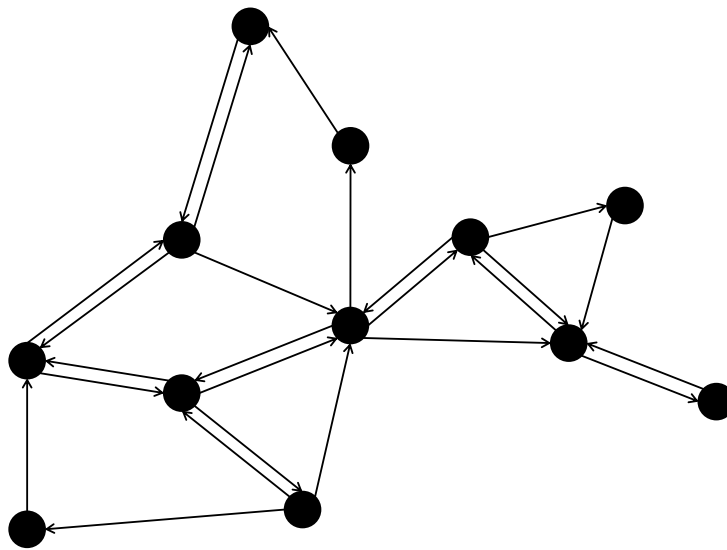


Abbildung 2.9: *Illustration eines gerichteten Graphen*

Generell werden im Graphen der digitalen Karte die Straßen durch Kanten und die Kreuzungen durch Knoten repräsentiert. Den Kanten sind Attribute, wie z.B. Länge und Geschwindigkeit oder eine Richtung, zugeordnet. In einer digitalen Karte können dies bis zu 200 sein [20]. Zusätzlich sind auch Informationen über den Zusammenhalt des Graphen abgelegt und beschreiben, welche Kanten in welchen Knoten aufeinandertreffen und damit *adjazent* zueinander sind. Eine Illustration einer Karte inklusive Attributen ist in Abbildung 2.10 dargestellt.

Die Erstellung einer digitalen Karte kann auf unterschiedlichen Wegen erfolgen. Die Anbieter für digitale Karten haben das zugrunde liegende Straßennetz mit speziellen Fahrzeugen abgefahren, es vermessen und Informationen erhoben, die als Kantenattribute in die Karte eingehen [20, 28]. Die einmal erhobenen Daten werden in Intervallen aktualisiert [28].

die Vorteile von OpenStreetMap zu nutzen und Nutzer an der Aktualisierung und Verbesserung von Karten zu beteiligen [28].

Das Ergebnis der Erhebung ist eine Karte für ein jeweiliges Gebiet, z.B. Deutschland oder Europa. In diesem Gebiet ist die Karte mit einem hohen Detailgrad gespeichert und kann den anderen Komponenten des Navigationssystems zur Verfügung gestellt werden. Ein vereinfachter schematischer Aufbau der Kartenspeicherung ist in Abbildung 2.12 dargestellt.

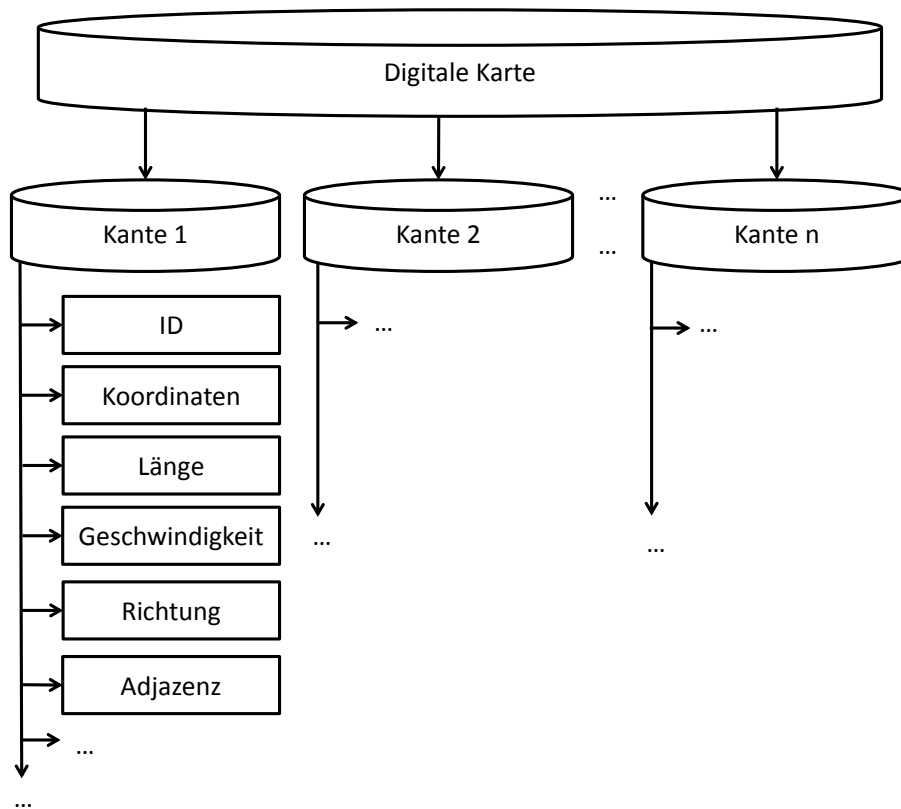


Abbildung 2.12: Vereinfachter Aufbau der digitalen Karte

Die Karte besteht aus einer Vielzahl von Kanten. Diese besitzen eine eindeutige Identität (ID) und die Koordinaten vom Anfangs- und Endpunkt der Kante. Dazu gibt es, wie oben bereits beschrieben, angehängte Attribute. Die wichtigsten sind dabei die Länge sowie die auf einer Kante erlaubte bzw. durchschnittlich gefahrene Geschwindigkeit. Weiterhin ist die Richtung, in die sie durchfahren werden darf (z.B. Einbahnstraßen), gespeichert. Dazu sind Informationen zu der oben beschriebenen Adjazenz zu anderen Kanten abgelegt.

2.2.2 Ortung

Bei der Ortung wird die aktuelle Position des Fahrzeugs bestimmt. Wie in Abschnitt 2.1.1 dargestellt, kann die Positionierung seit den 1990er Jahren ausschließlich satellitengestützt erfolgen. Am weitesten verbreitet ist dabei das amerikanische GPS. Dazu gibt es das russische System *GLONASS* sowie das sich im Aufbau befindende europäische *Galileo* [38]. Das Verfahren zur Ortung wird kurz beschrieben, ohne auf technische Details einzugehen. Dafür umkreisen Satelliten die Erde auf verschiedenen Umlaufbahnen (siehe Abbildung 2.13) und senden ein Signal mit der aktuellen Zeit sowie der Position aus [25]. Sowohl die Satelliten wie auch die Empfänger, z.B. im Navigationssystem, haben einen einheitlichen Zeitstandard. Aufgrund der Entfernung gibt es einen Verzug beim Satellitensignal, so dass daraus mit Hilfe komplexer Verfahren, z.B. Kalman Filter [99], der Abstand zu den Satelliten berechnet werden kann [92]. Da die Position der Satelliten bekannt ist, kann hierauf auf die Position des Fahrzeugs geschlossen werden.

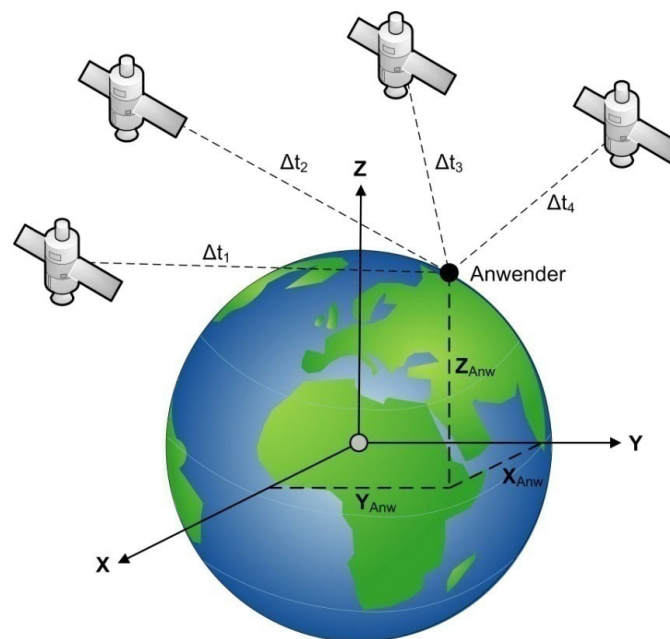


Abbildung 2.13: Schematische Darstellung der satellitengestützten Positionsbestimmung, Abbildung aus [92]

Für die Positionierung im dreidimensionalen Raum werden Signale von mindestens vier Satelliten benötigt. Diese sind verfügbar bei Sichtbarkeit der Satelliten. Dabei steigt die Qualität der Ortung mit der Anzahl der sichtbaren Satelliten. Weiterhin kann die Qualität durch Bodenstationen mit fester Position verbessert und eventuelle Störungen des Signals in der Atmosphäre ermittelt werden [92]. Für die

Berechnung der Position war ebenfalls die technische Entwicklung von entscheidender Bedeutung, da es nun auch für leistungsschwache Systeme möglich war, die Berechnungen in akzeptabler Zeit durchzuführen. Dies führte auch zur Entwicklung des *differentiellen GPS (DGPS)*, das die Ortung weiter präzisieren kann [58, 92].

Die größte Herausforderung bei der Ortung ist dabei der störungsfreie Empfang der Satellitensignale. Ist dies auf offenem Gelände, z.B. auf Autobahnen, gegeben, so ist der Empfang der Satellitensignale im urbanen Bereich deutlich schwieriger. Insbesondere hohe, dicht stehende Gebäude führen zur Blockierung oder Reflexion der Signale. In diesem Fall kann keine oder eine nur grobe Ortung erfolgen. In Tunneln ist dementsprechend eine Ortung unmöglich und die aktuelle Position wird auf Basis der letzten bekannten Position fortgeschrieben. In Abbildung 2.14 sind die verschiedenen Komponenten der Ortung sowie das anschließende Map Matching bzw. *Karteneinpassung* dargestellt.

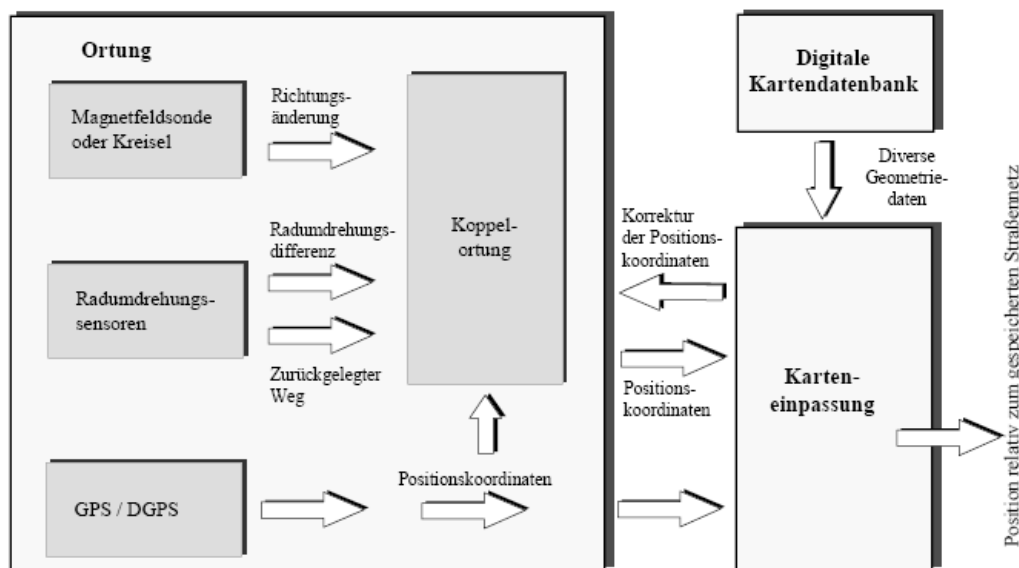


Abbildung 2.14: Aufbau des Ortungsverfahrens mit anschließendem Map Matching (Karteneinpassung), Abbildung aus [79]

Mit dem Fahrzeug verbundene Systeme können neben dem Satellitensignal auf Fahrzeugsysteme bzw. Sensoren zugreifen, um Drehung und damit die Ausrichtung des Fahrzeugs, aber auch die exakt zurückgelegte Wegstrecke für die Ortung mit einzubeziehen (*Koppelortung*) [79]. Unabhängig von der Güte der Position geht diese in den nächsten Schritt, das Map Matching, ein.

2.2.3 Map Matching

Beim Map Matching werden die Position der Ortung und die digitale Karte miteinander in Beziehung gesetzt und eine Position auf einer Kante der digitalen Karte bestimmt, siehe Abbildung 2.14, rechts (Karteneinpassung). Ziel ist es dabei, verschiedene Quellen und Verfahren zu nutzen, um eine möglichst korrekte Position auf einer Straßenkante zu ermitteln.

Dabei gibt es zwei Hauptfehlerquellen. Zum einen kann die Karte fehlerhaft erstellt (siehe Abschnitt 2.2.1) oder Straßen neu gebaut worden sein. Zum anderen kann es auch vorkommen, dass die Ortung fehlerhaft ist (siehe Abschnitt 2.2.2). Dazu werden die Straßen als Kanten bzw. Vektoren dargestellt und besitzen somit keine Breite. Da aber insbesondere mehrspurige Straßen eine große Breite besitzen, kann ein Fahrzeug sich an verschiedenen Positionen korrekt auf der Straße befinden. Somit stimmen die gemessene Position und die theoretisch korrekte Position häufig nicht vollkommen überein. In Abbildung 2.15 werden die Problematik der Eindimensionalität der Vektoren und der damit verbundenen Probleme beim Matching illustriert.

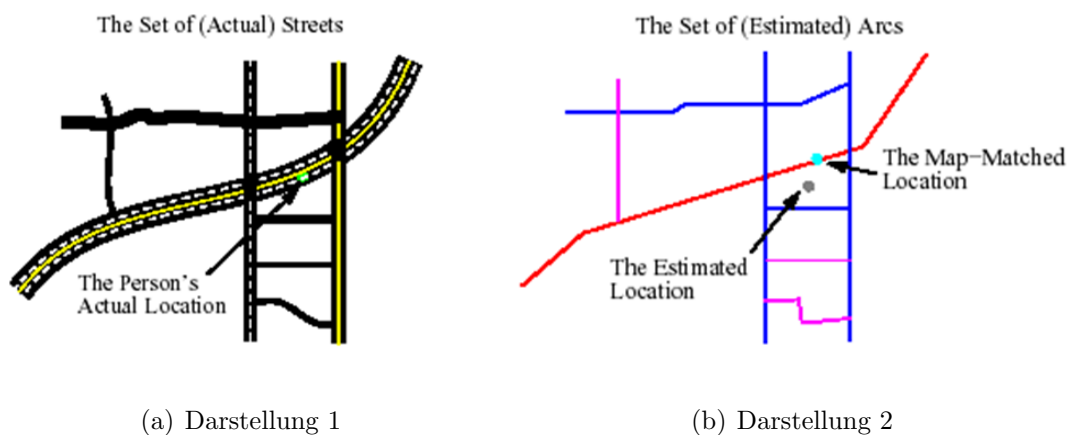


Abbildung 2.15: *Illustration des Map Matchings, Abbildung aus [43]*

Ein einfacher Map Matching Ansatz ist die Bestimmung des kürzesten Abstands der jeweiligen Position zu einer Kante und die Positionierung auf der Karte. Allerdings sind manche Straßen, insbesondere innerorts, häufig nahe beieinander, so dass es mit diesem rudimentären Ansatz zu fehlerhaften Map Matchings kommen kann. Dies wird noch dadurch verstärkt, dass es, wie bereits beschrieben, insbesondere im städtischen Bereich zu fehlerhaften Ortungen kommen kann.

In Navigationssystemen werden vielmehr mehrere Positionen zusammengefasst betrachtet. Die gemessenen Positionen unterliegen einer Struktur, die mit der Struktur des Straßennetzes abgeglichen wird. Somit lässt sich die Historie der Positionen im Zusammenspiel mit Informationen der digitalen Karte nutzen, um eine Trasse aus einer Auswahl möglicher Trassen zu bestimmen, auf der sich die korrekte Position befindet, siehe Abbildung 2.16 [43]. Die Verfahren zur Bestimmung der Trasse können variiert werden. Ziel ist es dabei immer, die wahrscheinlichste Trasse zu bestimmen, auf die sich der Fahrer mit seinem Fahrzeug bewegt. Dazu gibt es weitere inhaltliche Ansätze zur Verbesserung. [43, 79]

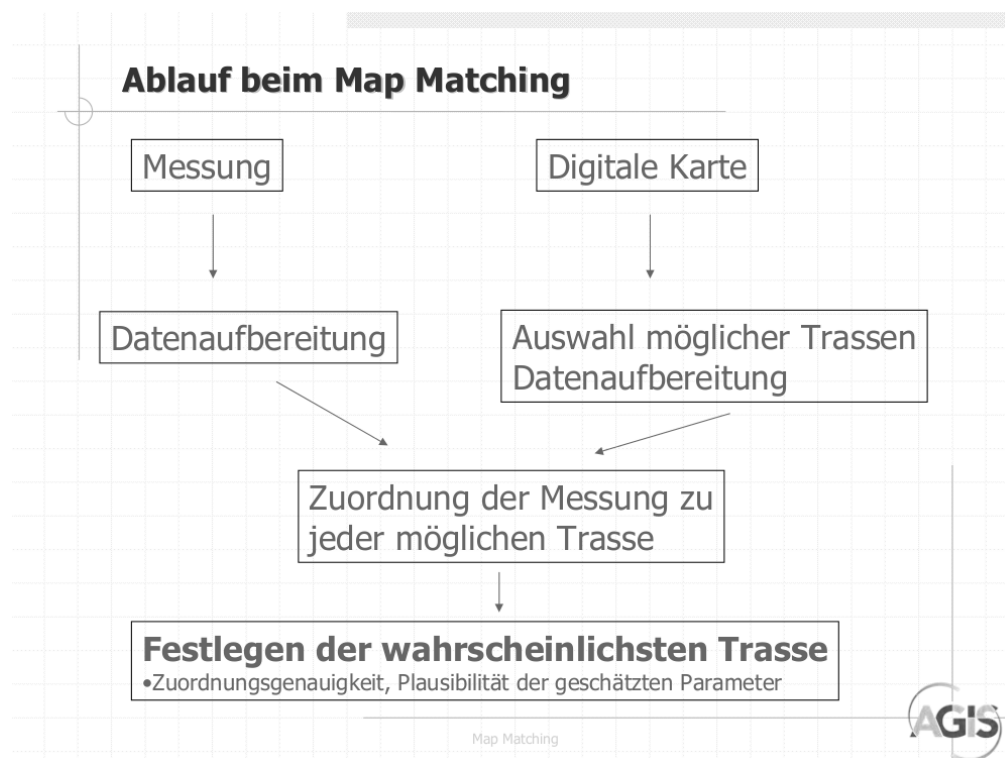


Abbildung 2.16: *Ablauf des trassenbasierten Map Matchings, Abbildung aus [79]*

Die einzelnen Verfahren und Ansätze zur Ermittlung der wahrscheinlichsten Trasse und damit der Position auf der Straße werden an dieser Stelle nicht im Detail beschrieben. Für detaillierte Darstellungen siehe bspw. [43, 79, 78]. In Abbildung 2.17 wird ein Beispiel illustriert, das mehrere Datenaufbereitungsschritte sowie Verfahrensansätze erfordert, um ein korrektes Map Matching zu ermöglichen.

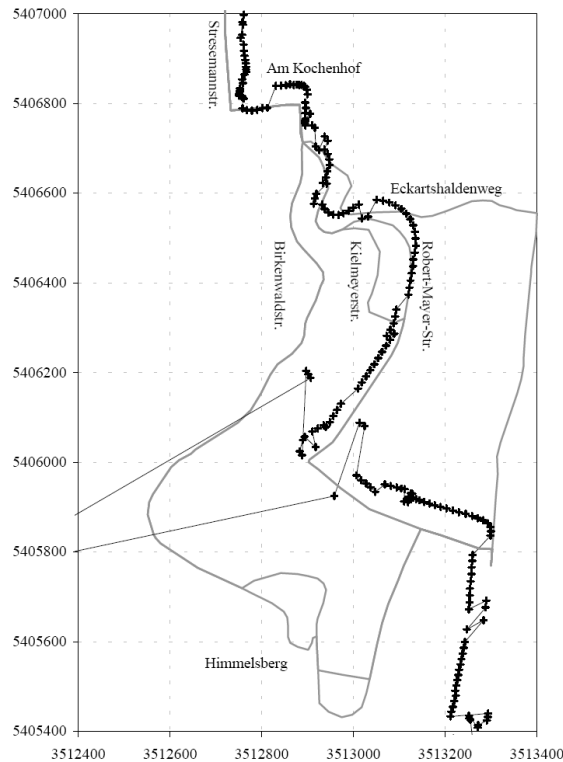
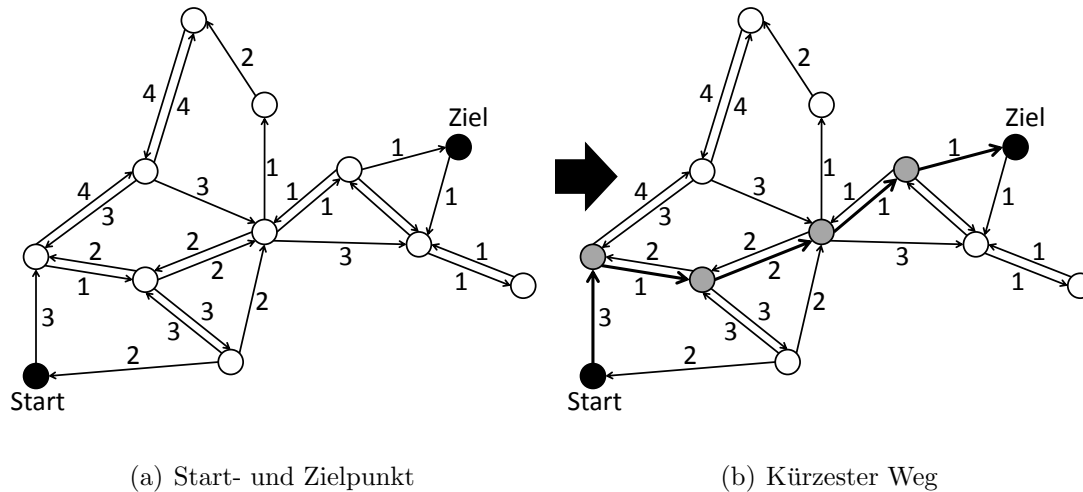


Abbildung 2.17: *Illustration der Herausforderungen des Map Matching Verfahrens, Abbildung aus [79]*

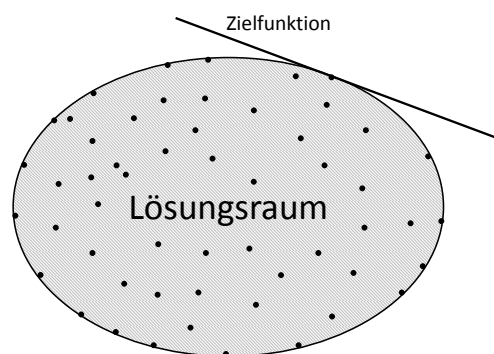
2.2.4 Routing

Die Routenberechnung bestimmt eine Route von einem Start- zu einem Zielpunkt. Sie ist der Kern eines Navigationssystems im Fahrzeug. Der Startpunkt ist im Allgemeinen die aktuell ermittelte Position. Der Zielpunkt wird dem Navigationssystem vom Fahrer mitgeteilt (siehe Abschnitt 2.2.5) und vom Navigationssystem einem Punkt auf dem Straßengraphen der digitalen Karte zugeordnet. Ebenso bestimmt der Fahrer ein Zielkriterium für die Route. Aufgabe der Routenberechnung ist es, einen Weg auf dem Graphen zu finden, um den Fahrer entsprechend der aus dem Zielkriterium abgeleiteten Zielfunktion optimal zum Ziel bringen zu können. Dabei werden KW-Verfahren [22, 55, 84] herangezogen, die einen solchen Weg berechnen. Eine Illustration eines solchen Weges ist in Abbildung 2.18 dargestellt.

Allen Verfahren zugrunde liegt die Minimierung eines Zielkriteriums. Meist ist dies eine schnellste bzw. kürzeste Route. Dementsprechend werden Gewichte gewählt, wie sie in Abbildung 2.18 dargestellt sind. Bei einer kürzesten Route werden die Kantenlängen als Gewicht genutzt. Bei einer schnellsten Route bekommen alle Kanten Zeitgewichte, berechnet aus der Kantenlänge und der in der digitalen Kar-

Abbildung 2.18: *Illustration eines kürzesten Weges*

te gespeicherten angenommenen Geschwindigkeit auf der jeweiligen Kante. Dieses Prinzip wird in allen Navigationssystemen angewandt und ist in Abbildung 2.19 vereinfacht illustriert. In der Abbildung sind alle möglichen Routen vom Start- zum Zielpunkt in einem Netzwerk als *Lösungsraum* dargestellt. Hierbei sei anzumerken, dass der Lösungsraum aus endlich vielen Lösungen bzw. Routen besteht, visualisiert durch die Punkte im Lösungsraum. In Abhängigkeit der Zielfunktion wird exakt eine Route ausgewählt. Der Unterschied zwischen der Route mit dem kleinsten sowie der Route mit dem nächstgrößeren Zielwert bezüglich der Zielfunktion kann allerdings beliebig klein sein.

Abbildung 2.19: *Lösungsraum der Routenberechnung*

Für die Bestimmung eines kürzesten Weges gibt es eine Vielzahl von exakten Verfahren, siehe Tabelle 2.20. Die *Komplexität* solcher Verfahren wird durch die maximale Laufzeit, dargestellt durch das *Landau-Symbol* O , bestimmt. Wichtigster

Faktor für die Laufzeit ist die Anzahl der Knoten V , sowie die Anzahl der Kanten E . Durch die Entwicklung verbesserter Verfahren konnte die Laufzeit verringert werden. Da es aber in einer digitalen Karte sehr viele Knoten gibt, sind diese klassischen Verfahren meist nicht für eine Routenberechnung, insbesondere für längere Routen, geeignet. Vielmehr werden praktikablere Verfahren, wie Heuristiken verwendet [86]. Diese erreichen eine in der Regel sehr gute Lösung innerhalb kürzester Zeit. Zusätzlich wird zur weiteren Verringerung der Laufzeit die digitale Karte in einer speziellen hierarchischen Struktur gespeichert, um zu vermeiden, dass das ganze Netzwerk durchsucht werden muss. Somit ist es möglich, nahezu jede Route innerhalb weniger Sekunden mit einer hohen Güte zu berechnen und dem Fahrer zur Verfügung zu stellen [86].

Algorithm	Time complexity	Author
	$O(V^2EL)$	Ford 1956
Bellman-Ford algorithm	$O(VE)$	Bellmann 1958, Moore 1959
	$O(V^2 \log V)$	Dantzig 1958, Dantzig 1960, Minty (cf. Pollack & Wiebenson 1960), Whiting & Hillier 1960
Dijkstra's algorithm with list	$O(V^2)$	Leyzorek et al. 1957, Dijkstra 1959
Dijkstra's algorithm with modified binary heap	$O((E + V) \log V)$	
...
Dijkstra's algorithm with Fibonacci heap	$O(E + V \log V)$	Fredman & Tarjan 1984, Fredman & Tarjan 1987
	$O(E \log \log L)$	Johnson 1982, Karlsson & Poblete 1983
Gabow's algorithm	$O(E \log_{E/V} L)$	Gabow 1983b, Gabow 1985b
	$O(E + W \log L)$	Ahuja et al. 1990

Abbildung 2.20: Übersicht gängiger Kürzeste-Wege-Verfahren, (V =Anzahl Knoten, E =Anzahl Kanten, L =maximales Gewicht einer Kante), Darstellung nach [84]

2.2.5 Human-Machine-Interface

Das HMI ist die Schnittstelle zwischen den internen Prozessen des Navigationssystems (Positionierung und Routenberechnung) sowie dem Fahrer. Zum einen ist es dem Fahrer möglich, Eingaben, wie z.B. die Zielposition, zu machen. Zum anderen soll der Fahrer vom System benutzerfreundlich zum Ziel geführt werden.

Damit wird der Nutzen eines Navigationssystems nicht nur über die Berechnung einer Route bestimmt sondern auch darüber, den Fahrer möglichst effizient und nutzerfreundlich an den Zielort zu führen.

Routenführung

Die Routenführung beinhaltet sowohl die graphische Aufbereitung der Route auf einer Karte wie auch die sprachliche Ausgabe von Anweisungen. Bei der Darstellung der Karte gibt es unterschiedliche Ansätze. Die Ausrichtung sowie der Detailgrad der Darstellung sind die wichtigsten Charakteristika. Dazu kommt die korrekte Darstellung der berechneten Route, um erfolgreich an dieser entlangfahren zu können [11]. In Abbildung 2.21 ist ein Beispiel dargestellt.

Die sprachlichen Ausgaben haben bei konkreten Manövern, wie Abbiegen, eine hohe Relevanz. Hier sind die korrekte Ausgabe sowie die zeitliche Abfolge von hoher Bedeutung. Die Ausgabe beinhaltet Abstandsangaben zum notwendigen Manöver und folgt häufig einer Eskalationsspirale. Für die zeitliche Anordnung dieser Befehle ist eine korrekte Ortung notwendig.



Abbildung 2.21: *Beispiel der Gestaltung der Route und des Menüs, Abbildung aus [66]*

Bedienung

Die Interaktion mit dem System findet beim Fahrer über die Bedienung statt. Meist erfolgt diese über Eingabe, es gibt aber auch Möglichkeiten zur Sprachsteuerung. Eingaben durch den Fahrer sind der Ausgangspunkt einer jeden Routenberechnung. Dazu gibt es dem Fahrer auch die Möglichkeit, während einer Fahrt mit dem System zu interagieren. Eine übersichtliche und selbst erklärende Menüführung ist wichtig. Neben dem inhaltlichen Aspekt ist auch die Designfrage von hoher Bedeutung, siehe Abbildung 2.21. Häufig sind dies die ersten Eindrücke, die der Fahrer

hat, noch bevor er vertraut ist mit dem System. Stärken und Schwächen einer Routenberechnung werden wiederum bei der Nutzung des Systems deutlich und prägen den Eindruck nachhaltig. Damit wird die Bewertung eines Navigationssystems größtenteils durch die Routenberechnung als wichtigste Komponente der Navigation bestimmt und hat somit das größte Potenzial für Ansätze zur Verbesserung des Navigationssystems.

2.3 Verbesserung der Routingkomponente der Navigation

2.3.1 Ansatzpunkte zur Verbesserung

Es gibt mehrere Ansatzpunkte zur Verbesserung der Routingkomponente von Navigationssystemen. Eine bessere Hardware kann zu einer schnelleren Berechnung oder einer detaillierteren Darstellung der Route führen. Ebenso kann durch verbesserte GPS-Empfänger die Ortung exakter werden. Dazu kann auf der Systemseite das Ortungs- oder das Matchingverfahren verbessert bzw. die Bedienung benutzerfreundlicher gestaltet werden. Nachhaltige Verbesserungen werden durch Ansätze, die direkt die Routingkomponente verändern, erreicht. Eine potentielle Möglichkeit ist es, einen neuartigen Algorithmus zu entwickeln, der nicht auf einem KW-Verfahren basiert. In Abschnitt 2.4.3 wird darauf näher eingegangen. Allerdings erfordert dies einen großen Aufwand in der Umsetzung sowie die Akzeptanz durch den Nutzer. In dieser Arbeit wird dagegen der Ansatz verfolgt, die Routenberechnung auf Basis bestehender Verfahren zu verbessern.

Die Routenberechnung basiert auf einer statischen digitalen Karte und einer festen Zielfunktion, die exakt eine Route bestimmt. Allerdings stimmen die Annahmen, die der Route zugrunde liegen, nicht immer mit der Realität überein. Es gibt zwei Faktoren, die hauptsächlich die Güte einer Route beeinflussen, siehe Abbildung 2.22. Dies sind zum einen externe Umweltfaktoren. Wettereinflüsse, wie z.B. Schnee, haben meist geringere Geschwindigkeiten zur Folge. Ein nahezu immer auftretender Einflussfaktor ist dazu der generell und aktuell herrschende Verkehr. Bei hohem Verkehrsaufkommen oder auch Störungen, wie z.B. Baustellen, kann die auf einer Kante angenommene Geschwindigkeit nicht erreicht werden.

Der zweite Faktor ist der Fahrer selbst. Jeder Fahrer fährt unterschiedlich, was sich in verschieden ausgeprägten Unterschieden zwischen der vom System angenommenen und generell vom Fahrer gefahrenen Geschwindigkeit äußern kann. Dazu kann

der Fahrer Anforderungen an eine Route sowie Präferenzen haben, die die berechnete Route nicht erfüllt. Dies kann die Bewertung einer Route durch den Fahrer negativ beeinflussen.

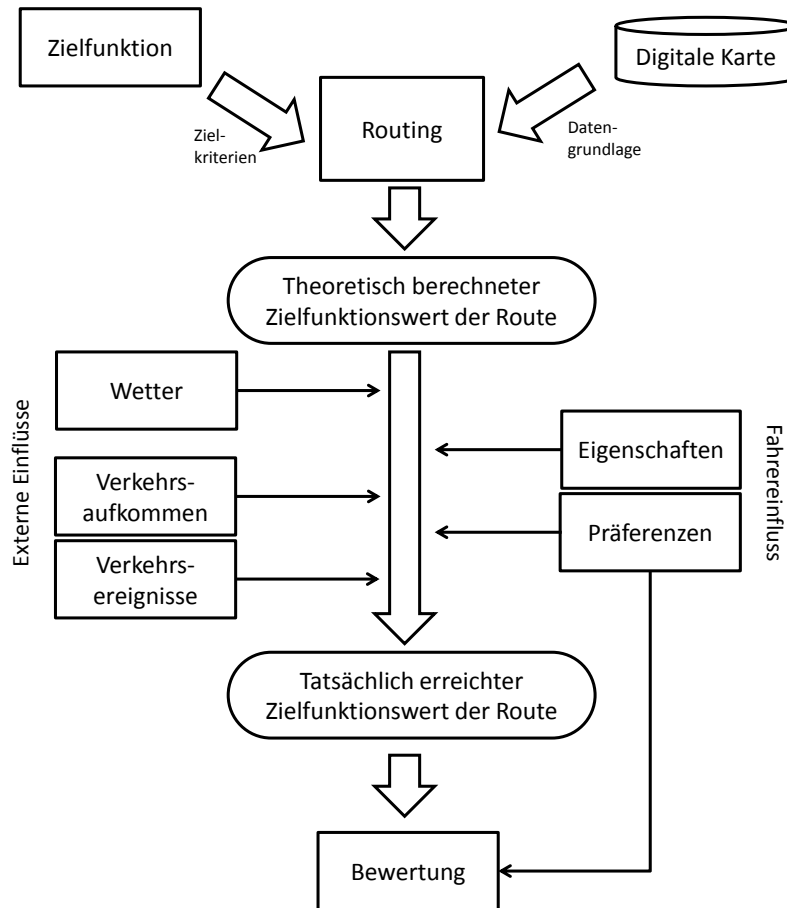


Abbildung 2.22: *Einflüsse auf den berechneten Zielfunktionswert der Routenberechnung*

Um diese Einflüsse für existierende Verfahren zur Routenberechnung und deren Auswirkungen zu berücksichtigen, gibt es verschiedene Möglichkeiten, die in Abbildung 2.23 visualisiert werden. So kann die Zielfunktion verändert werden bzw. eine vollkommen neue Zielfunktion für die Berechnung genutzt werden. Dadurch kann eine andere Route als die Route mit dem besten Zielfunktionswert bezüglich der neuen Zielfunktion ausgewählt werden (siehe Abbildung 2.23, links). Ebenso ist es möglich, durch Restriktionen Routen auszuschließen und so die ausgewählte Route zu verändern (siehe Abbildung 2.23, Mitte). Eine dritte Möglichkeit stellt die Veränderung der Kantengewichte dar (siehe Abbildung 2.23, rechts). In diesem Fall bleiben zwar alle Routen zulässig, doch kann es innerhalb des Lösungsraumes

zu Verschiebungen bezüglich des Zielkriteriums kommen. Es ist auch möglich, zwei oder auch alle drei Ansätze für eine Veränderung des Lösungsraumes und der Zielfunktion zu nutzen.

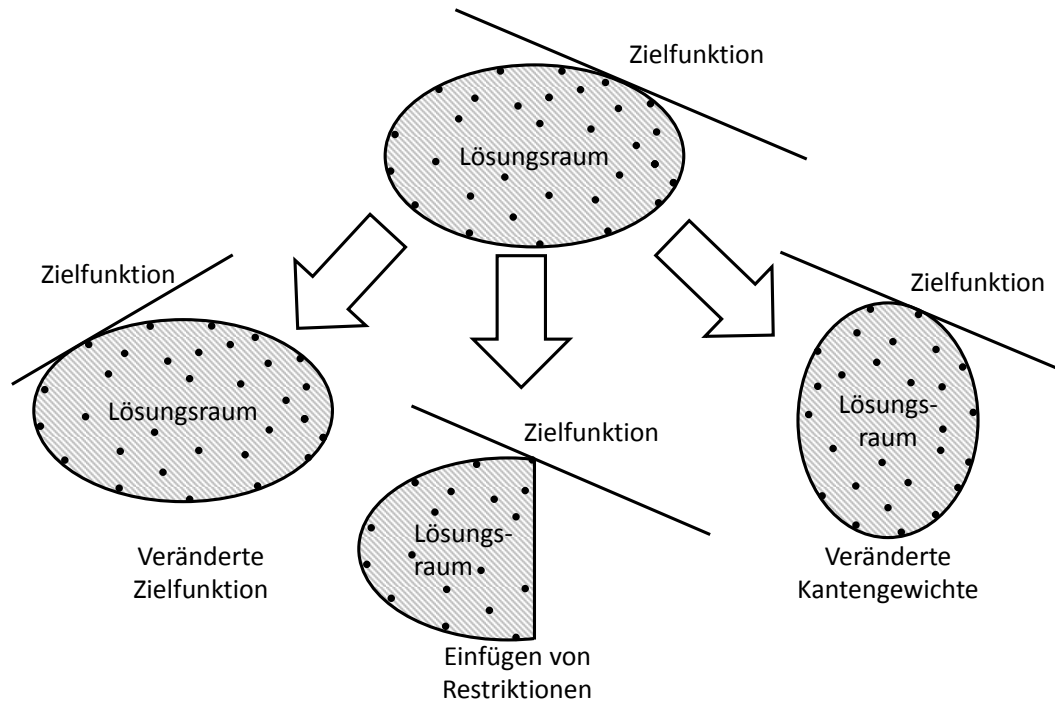


Abbildung 2.23: Visualisierung der Möglichkeiten zur Beeinflussung der Menge der zulässigen Routen sowie der optimalen Route

2.3.2 Einbeziehung externer Einflüsse

Ein wichtiger Faktor für die Güte einer berechneten Route ist der Verkehr. Dieser setzt sich zusammen aus dem generellen und dem aktuell herrschenden Verkehr. Generell gibt es für verschiedene Tages- und Wochenzeiten unterschiedliche Verkehrsaufkommen [7]. So ist das Verkehrsaufkommen Morgens und am späten Nachmittag aufgrund des Berufsverkehrs deutlich höher als bspw. nachts. Am Wochenende wiederum lassen sich diese erhöhten Aufkommen nicht finden. An diesen Tagen wird das Verkehrsaufkommen vornehmlich von Freizeitfahrten bestimmt.

Daher gibt es Ansätze für eine tageszeitabhängige Berechnung von Routen [24, 60, 71]. Je nach Tageszeit können unterschiedliche Geschwindigkeiten mit Hilfe von *Ganglinien* [7] als Kantengewichte für das Routing angenommen werden. Als Ergebnis können unterschiedliche Routen mit dem gleichen Zielkriterium für verschiedene Tageszeiten berechnet werden. Ebenso werden die Wochentage unter-

schieden und das Wochenende meist separat behandelt. Eine Visualisierung von Ganglinien ist in Abbildung 2.24 dargestellt. Dies entspricht dem rechten Fall aus Abbildung 2.23, dass der Lösungsraum verändert wird. Auf den Fahrer bezogen kann die Berücksichtigung von Ganglinien auch als Nachbildung des Verhaltens von erfahrenen, ortskundigen Fahrern, die ihre Routen je nach Tageszeit variieren und bestimmte verkehrsreiche Strecken meiden, angesehen werden.

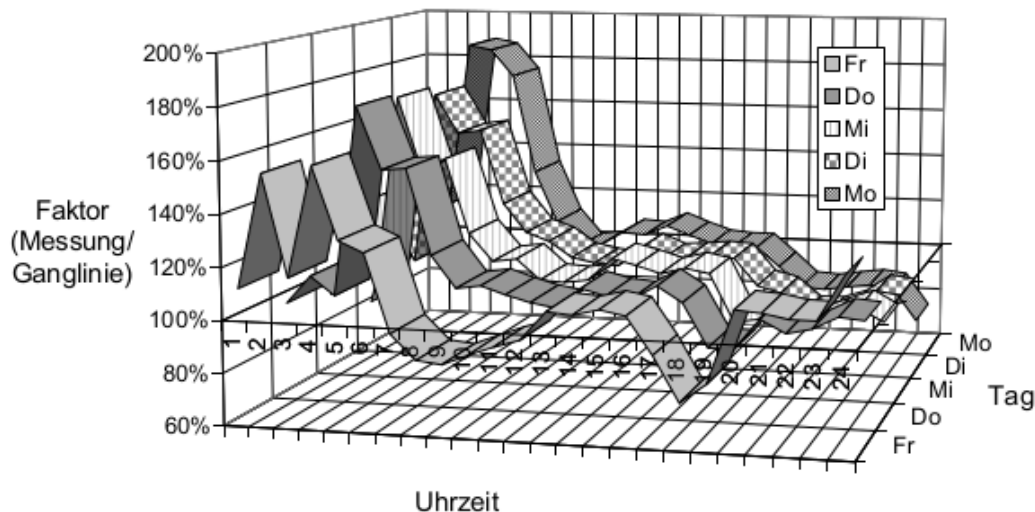


Abbildung 2.24: *Visualisierung von Ganglinien, Abbildung aus [7]*

Ähnlich verhält es sich, wenn es, z.B. aufgrund eines Unfalls, zu einer aktuell veränderten Verkehrslage kommt. Dies ist eine unvorhergesehene Veränderung, die nicht durch Ganglinien abgebildet werden kann. Erfahrene Fahrer könnten, wenn dies möglich ist, diese Strecke meiden und eine Alternativstrecke wählen. Ein Stau kann als eine Veränderung der Kantengewichte (siehe Abbildung 2.23, rechts), eine Vollsperrung als ein Hinzufügen einer Restriktion (siehe Abbildung 2.23, Mitte) angesehen werden. Grundlage dafür ist die Information über solch ein aktuelles Ereignis. Wenn der Fahrer durch eigene Anschauung solch ein Ereignis bemerkt, ist es für ihn meist nicht mehr möglich zu reagieren. Daher muss er für eine rechtzeitige Änderung der Route von anderen vorab darüber informiert werden.

Der klassische Weg hierfür sind dabei die amtlichen Verkehrsmeldungen. Diese werden im Radio vermeldet bzw. Navigationssystemen über den *Traffic Message Channel (TMC)* [23] zur Verfügung gestellt und können in die Routenberechnung durch Veränderung der entsprechenden Kantengewichte mit einbezogen werden. Das Ergebnis kann sein, dass eine Alternativstrecke empfohlen wird, aber auch dass

res Aufkommen und die Nutzung dieser Strecke würde zu einem besseren Ergebnis führen. In einem solchen Fall geht der Informationsvorsprung verloren.

Eine grafische Veranschaulichung der Verkehrslage aus dem Dienst von Google Maps ist in 2.26 illustriert. Diese ist frei zugänglich und kann die aktuelle Verkehrslage in die Routenberechnung mit einbeziehen. Allerdings ist anzumerken, dass somit sehr viele Nutzer Zugriff auf aktuelle Verkehrsdaten haben, was den oben genannten kontraproduktiven Einfluss potenzieren kann.

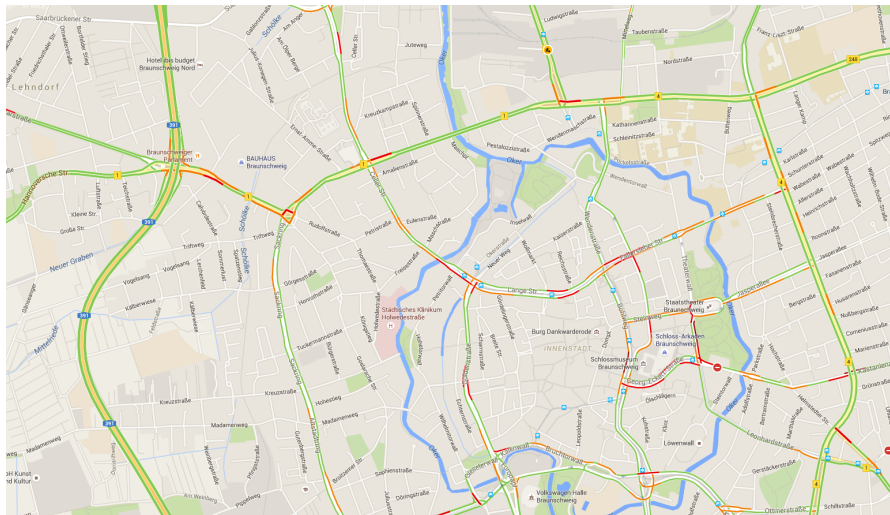


Abbildung 2.26: *Grafische Veranschaulichung des aktuell herrschenden Verkehrs aus [37]*

Um solche Fälle zu vermeiden, wird in der Verkehrssteuerung das Gesamtsystem betrachtet und versucht für das Gesamtsystem eine optimale Lösung zu finden, auch wenn es für den Einzelnen nicht optimal sein kann. Jeder Fahrer wird nur als Teil des Gesamtsystems angesehen [68, 88].

Dagegen hat ein System, wie oben beschrieben, den individuellen Nutzen eines Fahrers zum Ziel, allerdings wird er nicht als ein Individuum mit Eigenschaften und Präferenzen angesehen. Vielmehr wird eine Optimierung unter standardisierten Zielvorgaben unter Berücksichtigung der aktuell herrschenden Lage durchgeführt, unter der Annahme, dass dieses den Präferenzen des Fahrers entspricht.

Abschließend ist zu erwähnen, dass es mittlerweile auch möglich ist, das Wetter als externen Einfluss zu berücksichtigen. So bietet TomTom eine Funktion an, die eine wetterabhängige Route berechnet [23]. Dabei gehen Bedingungen wie Glatteis oder Starkregen, die die gefahrene Geschwindigkeit vermutlich verringern, in die Routenberechnung ein. Datenquellen sind aktuelle Wetterdaten, die einen vergleich-

baren Aktualitätsgrad haben wie die Verkehrsdaten. Auf diese Weise kann dieser Umwelteinfluss berücksichtigt werden.

Solche Wettereinflüsse wie auch Staus sind unvorhergesehen auftretende Ereignisse, die in eine aktuelle Routenberechnung mit einbezogen werden. Demgegenüber stellen Ganglinien, wie anfangs dargestellt, eine vorhergehende Anpassung an zu erwartende Umwelteinflüsse dar.

2.3.3 Einbeziehung des Fahrers

Der Fahrer ist ein weiterer wichtiger Einflussfaktor für die Güte einer Route. Zudem ist er derjenige, der das System nutzt und bewertet. Wenn der Fahrer das Navigationssystem nicht nutzt, wählt er selbstständig eine Route aus. Diese Kompetenz, eine Route von einem Start- zu einem Zielpunkt zu finden bzw. zu berechnen, gibt der Fahrer ab, wenn er selbst nicht über genügend Informationen verfügt, um zum Zielpunkt zu gelangen. Die Abgabe der *Navigationskompetenz* geschieht jedoch nicht, wenn der Fahrer die Aufgabe der Navigation selbst erfüllen will. So existieren zwei Systeme zur Routenberechnung bzw. Routenauswahl im Fahrzeug nebeneinander, siehe Abbildung 2.27.

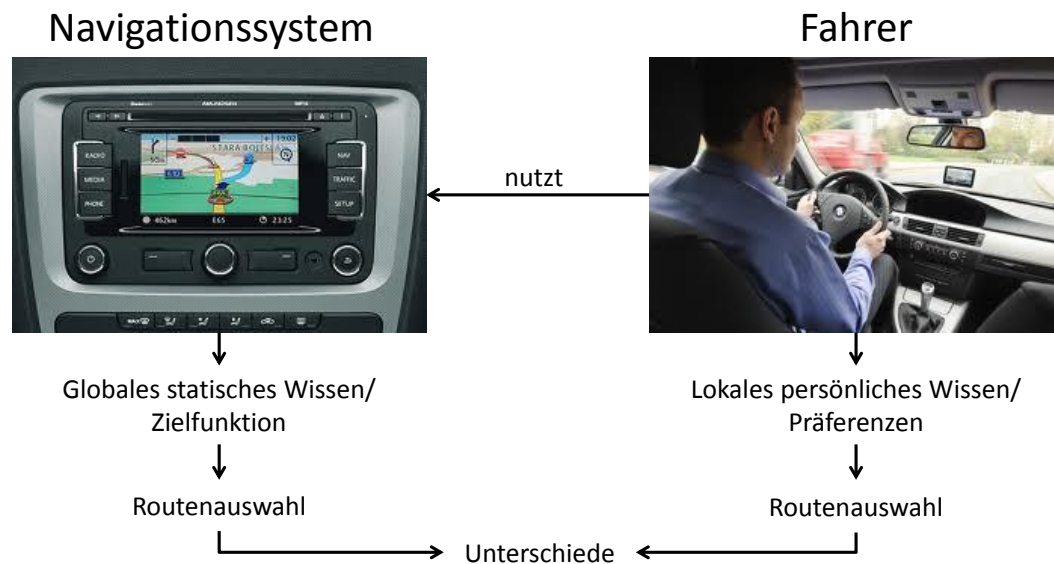


Abbildung 2.27: *Routenauswahlverhalten des Fahrers und des Navigationssystems*

Auf der einen Seite steht das Navigationssystem, das über globales aber meist statisches Wissen verfügt. Damit kann das System nahezu jede Route berechnen und

einen Fahrer dorthin führen, wenn der Zielpunkt in der digitalen Karte gespeichert ist. Die Routenberechnung geschieht dabei auf Basis dieses Wissens und der gegebenen Zielfunktion und Restriktionen.

Auf der anderen Seite steht der Fahrer mit einem lokalen persönlichen Wissen, das er ebenfalls für die Navigation einsetzt. Die Routenauswahl findet unter seinen Präferenzen und auf Basis des von ihm gesammelten Wissens über Straßen und Strecken, allerdings unter den genauen Vorgaben und Vorstellungen des Fahrers, statt. Somit kann es zu Unterschieden zwischen vom System und vom Fahrer ausgewählten Routen kommen. Hier sei anzumerken, dass das Wissen des Fahrers oftmals detaillierter und aktueller ist als das in der digitalen Karte abgelegte, so dass der Fahrer auf dem Gebiet, über das er das nötige Wissen verfügt, zu besseren Resultaten kommen kann als ein Navigationssystem.

Diese Unterschiede können den Nutzen des Navigationssystems für den Fahrer verringern, wenn sie nicht den Vorstellungen und Präferenzen des Fahrers entsprechen. Da der Fahrer nicht das gesamte Wissen der digitalen Karte erlernen kann, um das Routenauswahlverhalten des Systems nachzuvollziehen, muss eine Anpassung des Navigationssystems erfolgen. Eine solche Anpassung kann aktiv durch den Fahrer (*Adaptierung*) oder selbstständig durch das System (*Adaption*) erfolgen. Dadurch sollen die Unterschiede verringert und ein Mehrwert für das Navigationssystem generiert werden.

2.3.4 Adaptierung durch den Fahrer

Die Adaptierung von Navigationssystemen ist ein weit verbreiteter Ansatz und in nahezu jedem Navigationssystem verfügbar. Wie bereits in Abbildung 2.23 dargestellt, gibt es die Möglichkeiten sowohl die Zielfunktion als auch die Datengrundlage für die Berechnung anzupassen. Die am Häufigsten gewählte Zielfunktion betrifft die Dauer einer Route (schnellste Route). Ebenso ist die kürzeste Route, die eine Route mit minimaler Streckenlänge berechnet, verbreitet. Von diesen klassischen Zielkriterien abgeleitet ist eine *optimale Route*. Hierfür werden sowohl die Dauer wie auch die Länge berücksichtigt. Dabei können sie unterschiedlich stark gewichtet und vom Fahrer eingestellt werden. Ein Beispiel für ein weiteres Kriterium wäre eine *ökologische Route*, die den Kraftstoffverbrauch einer Fahrt minimiert. Auch Kriterien wie Kosten aber auch Komfort sind als maßgebliche Kriterien denkbar und werden in Ansätzen verfolgt [26]. Eine Visualisierung der Auswirkungen einer Veränderung der Zielfunktion war in Abbildung 2.23, links dargestellt. In diesem

Fall kann es dazu kommen, dass eine andere optimale Route aus der Menge der möglichen Routen ausgewählt wird.

Ebenso kann der Fahrer Restriktionen einfügen bzw. die Kantengewichte verändern (siehe Abbildung 2.23, Mitte und rechts). Ein Beispiel ist die Einstellung für die Autobahnbenutzung. Grundsätzlich gibt es drei Einstellungen: „Autobahn erlaubt“, „Autobahn vermeiden“ und „Autobahn verboten“. Ist die Benutzung von Autobahnen erlaubt, wird keinerlei Veränderung an den Daten vorgenommen und Autobahnabschnitte werden so gewichtet, wie sie ihren Attributen laut der digitalen Karte entsprechen. Wird die Einstellung „Autobahn vermeiden“ gewählt, so werden die Gewichte wie die angenommene Geschwindigkeit der entsprechenden Kanten auf der Autobahn verändert und der Lösungsraum angepasst. So werden Alternativmöglichkeiten aufgewertet. Ist der Effekt der negativen Bewertung größer, so steigt die Wahrscheinlichkeit, dass eine andere Route bevorzugt wird bzw. Teile der Autobahnstrecke ersetzt werden.

Dieser Effekt lässt sich erhöhen, indem sämtliche Autobahnen nicht berücksichtigt werden, wenn die Option „Autobahn verboten“ benutzt wird. In diesem Fall werden Restriktionen eingeführt und alle Routen mit Autobahnnutzung als nicht zulässig klassifiziert. Auch andere Ereignisse wie Mautstellen oder Fähren können so vermieden werden. Dazu werden gezielt einzelne Kanten ausgeschlossen. Hier sei anzumerken, dass dies allerdings dazu führen kann, dass entweder kein Weg zum Ziel mehr berechnet werden kann oder sich die angebotene Route deutlich verschlechtert. Dies steht oftmals in keinem Verhältnis zum angegebenen Wunsch etwas zu vermeiden.

Zusätzlich ist es dem Fahrer in vielen Navigationssystemen möglich, nach seiner Einschätzung passende Geschwindigkeiten auf unterschiedlichen Straßenklassen wie Autobahnen oder innerorts auszuwählen und so die Kantengewichte in seinem Sinne zu verändern. Waren die vorher beschriebenen Einstellungen noch auf bestimmte Punkte oder Klassen begrenzt, so hat die Einstellung der Geschwindigkeiten Auswirkungen auf die Bewertung der gesamten in der digitalen Karte gespeicherten Straßen bzw. Kanten. Insbesondere bei Auswahl der schnellsten Route hat dies einen großen Einfluss.

Dies zeigt, dass der Fahrer durch diese wie auch andere Funktionen Einfluss auf die Routenberechnung nehmen kann, mit dem Ziel, dass die vom Navigationssystem berechnete Route seinen Anforderungen und Präferenzen entspricht. Durch diese manuell durchgeführte Anpassung sollen die Unterschiede zwischen der vom System ausgewählten Route und der vom Fahrer ausgewählten verringert werden. Dies

ist in Abbildung 2.28 dargestellt und lehnt sich an die Darstellung aus Abbildung 2.27 an.

Allerdings ist die Adaptierung mit Aufwand seitens des Fahrers verbunden. Gerade im Straßenverkehr sollte aber dieser zusätzliche Aufwand vermieden werden [26]. Ebenso kann es zu Fehleinschätzungen durch den Fahrer z.B. bei der gefahrenen Geschwindigkeit kommen. Meist werden auch die eigenen Fahrfähigkeiten überschätzt. Bei der Routenberechnung kann eine Falschadaptierung zu unerwünschten Ergebnissen führen und den Nutzen deutlich verringern.

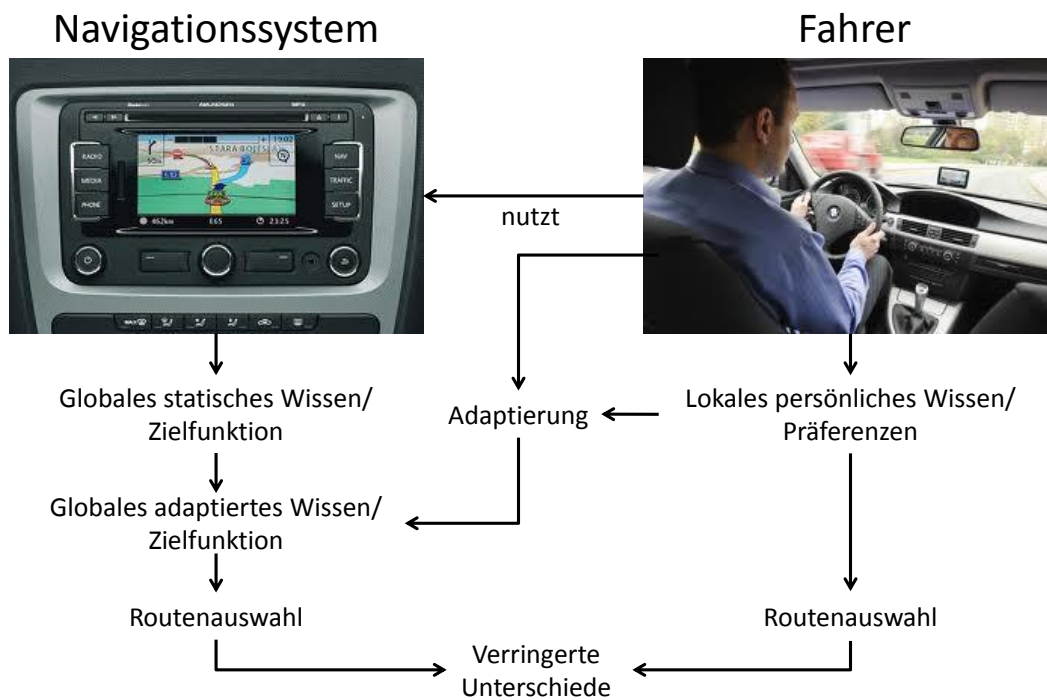


Abbildung 2.28: Adaptierung des Navigationssystems durch den Fahrer

2.3.5 Personalisierung durch Adaption an den Fahrer

Bei einer Adaption des Systems an den Fahrer bzw. Nutzer ist im Gegensatz zum adaptierbaren System kein aktives Handeln des Nutzers notwendig. In dieser Arbeit wird die selbstständige Anpassung des Navigationssystems an den Fahrer als Personalisierung bezeichnet. Ein Beispiel für eine solche Personalisierung ist der Ansatz für adaptive Geschwindigkeiten, die für die Routenberechnung genutzt werden.

Diese Funktion der adaptiven Geschwindigkeit muss zwar noch vom Fahrer manuell ausgewählt werden, doch funktioniert die Anpassung danach vollkommen selbstständig und passt sich immer weiter an. Es werden die vom Fahrer gefahre-

nen Geschwindigkeiten auf den einzelnen Straßenklassen gemessen und zu einem Profil der Geschwindigkeiten auf den einzelnen Klassen aggregiert. Dieser Ansatz kann beliebig komplex sein. Am praktikabelsten und schnell zu realisieren ist dabei ein Konzept des Mittelwertes. Dieser ergibt generell ein qualitativ ausreichendes Bild und benötigt nur wenig Speicher- und Rechenaufwand. Daher ist der gewählte Ansatz ein erster wichtiger Schritt, das Navigationssystem oder zumindest Teile davon zu personalisieren.

Das Vorgehen bei einem adaptiven System ist in Abbildung 2.29 dargestellt. Das Navigationssystem nutzt die vom Fahrer selbst ausgewählten Routen bzw. das durch seine Präferenzen bestimmte Verhalten für die Analysen zur Adaption. Wird eine Route mit Hilfe des Navigationssystems berechnet, findet dies auf Basis des globalen aber adaptierten Wissens statt mit dem Ziel, die Unterschiede zu verringern.

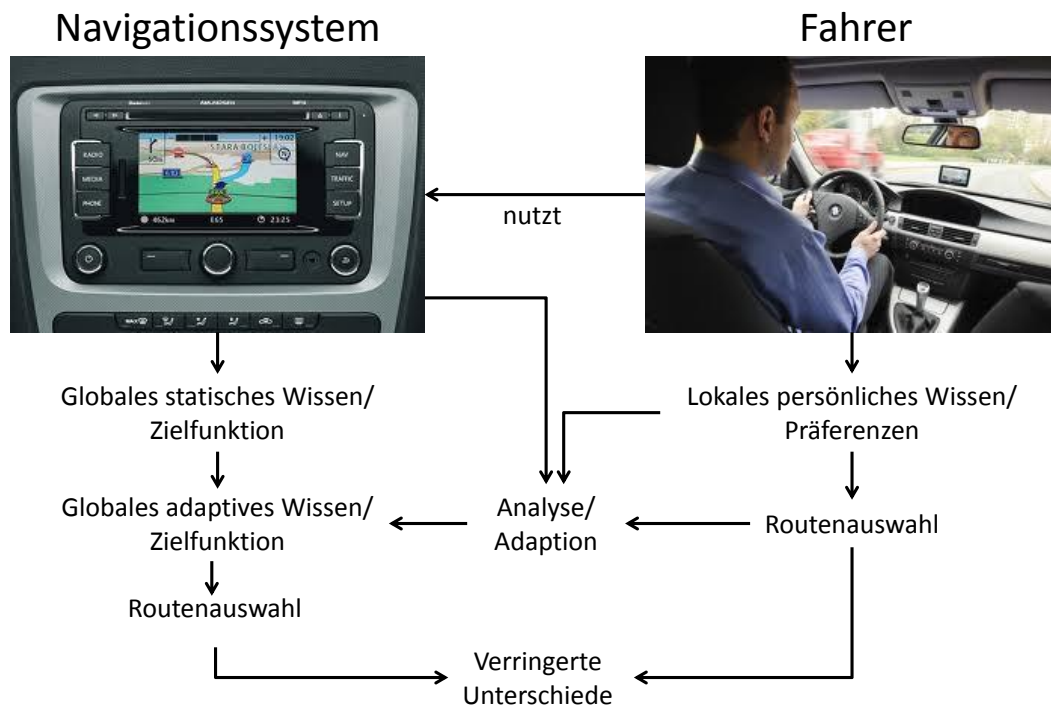


Abbildung 2.29: *Adaption des Navigationssystems an den Fahrer*

Grundlage einer jeden Personalisierung ist ein Modell des Fahrers. Bei der adaptiven Geschwindigkeit ist dies ein, wenn auch relativ einfaches, Modell über das Geschwindigkeitsverhalten des Fahrers. Die Qualität hängt dabei zuallererst von der Menge an Daten ab, die für die Bestimmung zur Verfügung stehen. Je größer die Menge ist, desto gesicherter lassen sich Aussagen treffen und diese für die Rou-

tenberechnung nutzen. Größere Auswirkung auf die vom Fahrer wahrgenommene Güte hat auch die Streuung der Daten. Insbesondere durch Verkehr aber auch Änderungen des Fahrstils können die Geschwindigkeiten stark variieren. Dies kann bei standardisierten Werten dazu führen, dass die Routenberechnung auf falschen bzw. ungenauen Annahmen aufgebaut wird.

Um dieses zu berücksichtigen ist ein Modell des Fahrers notwendig, das nicht ausschließlich auf seinem Geschwindigkeitsverhalten basiert, sondern mehrere Komponenten umfasst. In dieser Arbeit wird ein solches Modell entwickelt und angewandt. Der theoretische Prozeß ist in Abbildung 2.30 dargestellt. Daten, die über den Fahrer erhoben wurden, müssen analysiert und in Wissen umgewandelt werden. Das so generierte Wissen wird in ein Modell des Fahrers überführt. Dieses Modell kann genutzt werden, um eine Personalisierung durchzuführen. Der Prozess verläuft analog zum bekanntem KDP, siehe Abschnitt 3.3.

Dem Fahrer Eigenschaften zuzuordnen und seine Präferenzen zu bestimmen, um diese in ein Modell einfließen zu lassen, ist auch oftmals Ziel in der Verkehrspsychologie [16, 26]. Bei der Bewertung vorhandener Ansätze in Kapitel 2.4 wird auf die Modellierung näher eingegangen. An dieser Stelle ist festzuhalten, dass Personalisierung als selbstständiger Prozess, wie sie hier verstanden wird, im Bereich der Navigation bisher nur sehr rudimentär stattfindet. Dem Nutzer ist es oft nur möglich, Einstellungen manuell vorzunehmen und so das System entsprechend seinen Vorstellungen zu adaptieren.

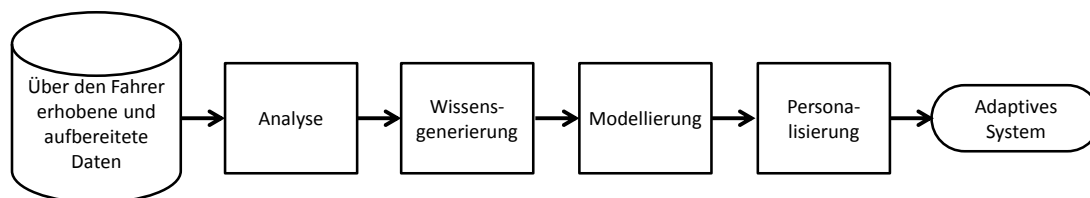


Abbildung 2.30: *Prozess der Personalisierung*

2.4 Analyse existierender Personalisierungsansätze und -verfahren

Da die Personalisierung der Routingkomponente von Navigationssystemen bisher nur in Ansätzen erfolgt, werden in diesem Abschnitt verschiedene existierende und zum Teil modellgestützte Verfahren zur Personalisierung aus unterschiedlichen Ge-

bieten evaluiert. Zunächst werden weit entwickelte Verfahren zur Personalisierung aus anderen Bereichen dargestellt und ihre Übertragbarkeit analysiert. Daraufhin wird untersucht, ob und welche Ansätze es zur Anpassung an den Fahrer im Fahrzeugbereich bereits gibt und erfolgreich eingesetzt werden. Abschließend werden Forschungsansätze im Bereich der Navigation erläutert sowie Stärken und Schwächen dargestellt und insbesondere die Nutzbarkeit diskutiert. Anhand der dargestellten Arbeiten werden die Anforderungen an eine Personalisierung der Routingkomponente eingeführt und erläutert.

2.4.1 Verfahren in nicht-fahrzeugbezogenen Gebieten

Da eine erfolgreiche Personalisierung von Systemen eine hervorragende Möglichkeit bietet, diese an den jeweiligen Nutzer anzupassen und so den Wert für den Benutzer erheblich zu steigern, ist sie in manchen Gebieten weit verbreitet. Dabei kann eine Personalisierung sowohl bei den Funktionen des Systems als auch bei der Interaktion mit dem Nutzer stattfinden.

Eine Vorreiterrolle im Bereich selbstständiger Personalisierung spielen internetbasierte Anwendungen der *Informationstechnologie (IT)*. Hier wurde schon früh damit begonnen, Systeme an den Nutzer anzupassen. Dies betrifft sowohl die Gestaltung der Benutzeroberfläche wie auch das Verhalten des Systems selbst. Dabei stellt die große Zahl von Nutzern den größten Vorteil dar. Dadurch hat man einen sehr großen Datenpool zur Verfügung, in dem Klassifizierungen oder Zuordnungen von Charakteristika durchgeführt werden können. Im *E-Commerce*, wie z.B. beim Internethändler Amazon, haben solche Verfahren mittlerweile einen hohen Entwicklungsstand erreicht.

Als Methode eignet sich insbesondere das sogenannte *Collaborative Filtering* [81]. Dabei werden Daten über die verschiedenen Nutzer gesammelt, um Verhaltensmuster von Benutzergruppen zu erkennen und daraus auf die Interessen des Einzelnen zu schließen. Dieser wird einer Benutzergruppe oder Klasse zugeordnet bzw. Nutzer zu einer Klasse zusammengefasst, siehe Abbildung 2.31. [81, 89].

Darauf aufbauend kann eine Ähnlichkeitsmatrix [87] zwischen den Nutzern eines Angebots gebildet werden, siehe Abbildung 2.32. Damit können den Nutzern jene Elemente präsentiert werden, die von ihren statistischen Nachbarn genutzt und/oder positiv bewertet wurden und damit eine Personalisierung des Systems erfolgen, was den Vorteil bietet, dass einem Nutzer ein angepasstes Angebot gemacht werden kann, wenn er einer Gruppe zugeordnet wurde.

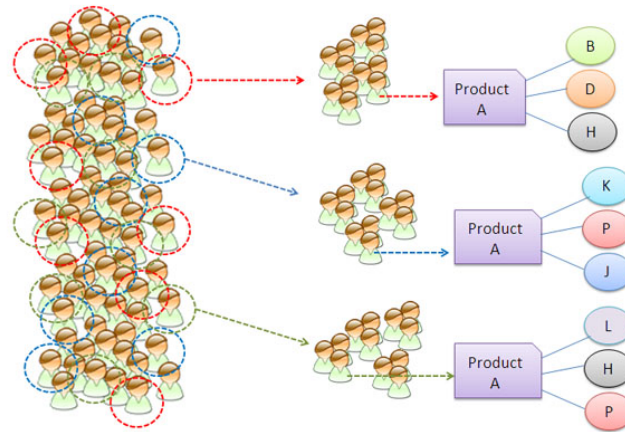


Abbildung 2.31: Klassifizierungsprozess beim Collaborative Filtering, Abbildung aus [89]

Ein spezifisches Problem kollaborativer Filter besteht allerdings in ihrer Latenzzeit [81]. Ein neuer Nutzer tritt mit einem leeren Benutzerprofil in das System ein und kann somit zu Beginn keine sinnvollen Empfehlungen erhalten. Gleiches gilt für neu in das System eintretende Elemente (z. B. Produkte in einem Online-Shop). Sie weisen keine quantifizierbare Ähnlichkeit mit anderen Elementen auf und können damit nicht sinnvoll empfohlen werden. Es handelt sich bei kollaborativen Filtern also um lernende Systeme auf einer großen Datenbasis unterschiedlicher Nutzer.

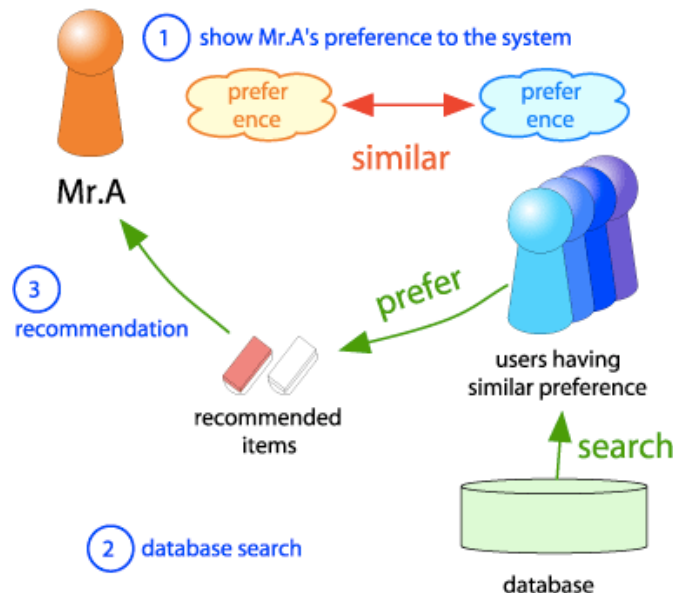


Abbildung 2.32: Zuordnungsprozess beim Collaborative Filtering, Abbildung aus [89]

Grundlage aller Verfahren ist der Zugang zu erhobenen Daten von verschiedenen Nutzern. Dies steht im engen Zusammenhang zur Verfügung und Nutzung der Daten. Während im Internet persönliche Daten und Verhalten wissentlich oder unwissentlich weitergegeben werden bzw. zugänglich sind, ist die Hemmschwelle zur Weitergabe von Daten über das Verhalten im Verkehr deutlich höher. Die Folgen sind dort für den Nutzer direkt ersichtlich. Ein Überschreiten der Geschwindigkeitsbegrenzung ist eine reale Handlung und könnte bei einem Übermitteln der dazugehörigen Daten über die Geschwindigkeit geahndet werden. Daher ist, neben technischen Einschränkungen, nicht zu erwarten, dass ein großer Datensatz von Fahrerdaten verfügbar ist, um darauf Collaborative Filtering Verfahren darauf anzuwenden. Vielmehr sind Verfahren notwendig, die ausschließlich mit von einem Fahrer erhobenen Daten arbeiten.

Im IT-Bereich ist diese Situation im *E-Learning* [42] gegeben. Unter E-Learning werden alle Formen von Lernen verstanden, bei denen elektronische oder digitale Medien für die Präsentation und Distribution von Lernmaterialien und/oder zur Unterstützung zwischenmenschlicher Kommunikation zum Einsatz kommen. Dies bietet weitreichende Möglichkeiten die Lernfunktionen anzupassen im Vergleich zu herkömmlichen Lernmitteln wie einem Buch, das statisch ist. Hier soll durch die Personalisierung einen Nutzen für den Nutzer generiert werden. Durch angepasstes Lehrverhalten an das Lernverhalten soll der Nutzer weder unter- noch überfordert und die Lernkurve gesteigert werden.

Für die Akzeptanz bietet sich der Vorteil, dass nur das System und der Nutzer entscheidend sind und ein geschlossenes System bilden, siehe Abbildung 2.33. Die Erfolgsmessung kann direkt am Verhalten des Nutzers abgelesen und anhand von vorher festgelegten Kriterien eingeordnet werden. Die Einbindung oder Weitergabe sensibler Daten an die Außenwelt ist dabei nicht erforderlich. Hieran zeigt sich auch einer der wichtigsten Unterschiede zwischen dem Verhalten am Computer und dem im Fahrzeug. Beim Fahren eines Fahrzeuges spielt die Umwelt eine sehr wichtige Rolle, während die Benutzung eines Computers nur peripher von äußeren Einflüssen beeinflusst wird.

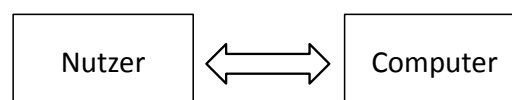


Abbildung 2.33: *Schaubild des Systems Nutzer-Computer*

Im Gegensatz dazu ist der Fahrer mit seinem Fahrzeug, wie in Abschnitt 2.3.2 beschrieben, dauernd der Umwelt ausgesetzt, siehe Abbildung 2.34. Diese wiederum verändert sich fortlaufend und erfordert eine Anpassung oder zumindest eine Berücksichtigung bei der Auswertung des Verhaltens. Dazu gibt es einen stetigen Austausch von Informationen. Daher können im E-Learning entwickelte Verfahren nicht übernommen werden, sondern bedürfen zumindest einer Anpassung an die besondere Situation im Fahrzeug.

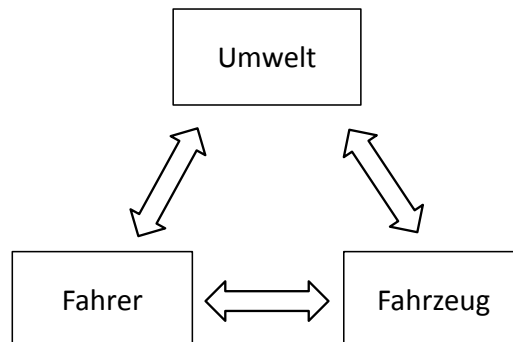


Abbildung 2.34: *Schaubild des Systems Fahrer-Fahrzeug-Umwelt*

2.4.2 Implementierte Ansätze im Fahrzeug

Im Fahrzeugbereich gibt es bereits Ansätze für Personalisierungen insbesondere für *Fahrerassistenzsysteme (FAS)*. Sie basieren auf den Daten eines Fahrers. Auch hier ist die Umwelt meist ebenso relevant wie für Navigationssysteme und alle Komponenten des Dreiecks aus Fahrer-Fahrzeug-Umwelt müssen berücksichtigt werden. Personalisierungsansätze gibt es für eingreifende aber auch für warnende Systeme. Hierbei wird im zunehmenden Maße versucht, die Assistenzsysteme an den jeweiligen Fahrer anzupassen. Jeder Fahrer hat einen eigenen Fahrstil und bewertet Nutzen aber auch Störung durch das FAS von seiner Sicht- und Fahrweise aus. Daher ist es von großem Vorteil für ein FAS ein entsprechendes Modell vom Fahrtyp des Fahrers zu bilden [48]. Dieses Modell kann zur Anpassung der Funktionen genutzt werden, um so den Nutzen aber auch die Nutzungshäufigkeit des FAS und damit auch Sicherheit und Komfort für den Fahrer zu erhöhen. Damit wird das Assistenzsystem Teil der Interaktion zwischen Fahrer, Fahrzeug und Umwelt, siehe Abbildung 2.35.

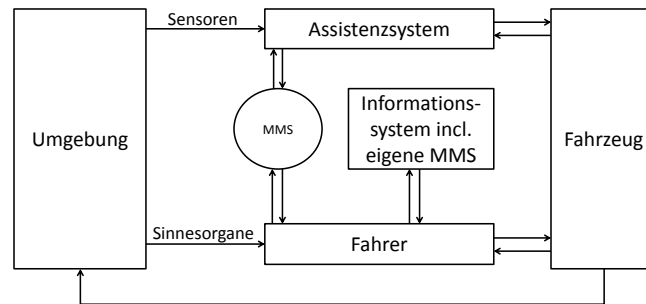
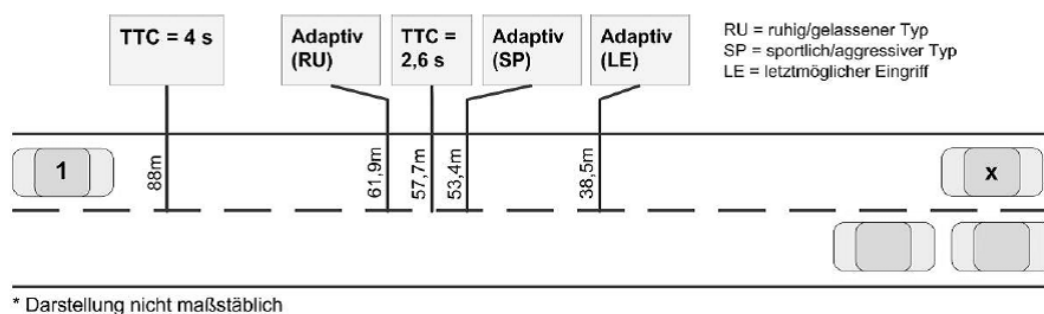


Abbildung 2.35: Schaubild des Systems Fahrer-Fahrzeug-Umwelt für FAS, Abbildung nach [48]

Ein Beispiel ist ein adaptives *Bremsassistentensystem*. Ein standardmäßig eingestelltes System kann bei einem sportlichen Fahrer, der einen Bremsvorgang relativ spät einleitet, zu restriktiv eingestellt sein. Dadurch wird der Fahrer oftmals unnötigerweise gewarnt, da die Annäherung an ein Hindernis erfolgt, er aber aufgrund seines Fahrweise noch nicht mit dem Bremsvorgang begonnen hat. Die Folge der wiederholt auftretenden unnötigen Warnungen kann ein Abschalten des FAS sein, das damit keinen Nutzen mehr hat und in einer relevanten Situation nicht mehr warnen bzw. eingreifen kann [85]. Hier, wie auch bei anderen Systemen, wäre eine Anpassung an den jeweiligen Fahrtyp von großem Nutzen. In Abbildung 2.36 ist eine solche Anpassung dargestellt. Es werden zwei Fahrtypen (RU, SP) und der letztmögliche Eingriffszeitpunkt (LE) unterschieden und unterschiedliche Schwellen für den Bremsassistenten gesetzt. [85]



* Darstellung nicht maßstäblich

Abbildung 2.36: Darstellung eines adaptiven Bremsassistenten, Abbildung aus [85]

Zur Einordnung solcher Ansätze wird die Fahraufgabe insgesamt betrachtet. In der Verkehrspsychologie wird diese klassisch in drei Ebenen unterteilt, siehe Abbildung 2.37 [8]. Die untere Ebene ist die *Stabilisierungsebene*. Diese umfasst alle

Aktivitäten, die der Fahrer durchführt, um das Fahrzeug auf der Straße zu stabilisieren. Es umfasst Geschwindigkeitsanpassungen genauso wie Lenkvorgänge. Die mittlere *Bahnführungsebene* umfasst ganze Manöver, wie z.B. Rechts-Abbiegen oder Spurwechsel. Der Fahrer muss diesen einleiten und erfolgreich durchführen.

Die oberste Ebene ist die *Navigationsebene*. Auf dieser trifft der Fahrer Entscheidungen, die die Route betreffen. So zum Beispiel die Wahl der Straße wie auch Entscheidungen, ob abgebogen werden soll, wenn es die gewählte Route verlangt. Mit steigender Ebene sinkt zwar die Häufigkeit mit der Aktionen auf der entsprechenden Ebene durchzuführen sind, dafür steigen aber die kognitiven Anforderungen. Im Fahrzeug genutzte FAS, statisch oder adaptiv, betreffen die beiden unteren Ebenen. Daher sind Modellierungs- und Personalisierungsansätze auf diesen Ebenen nur bedingt nutzbar für die Navigationsebene.

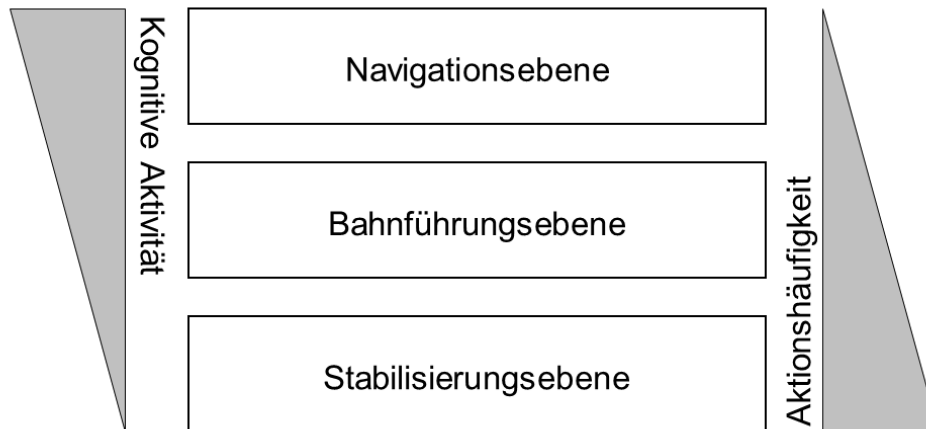


Abbildung 2.37: Darstellung des 3-Ebenenmodells der Fahraufgabe, Abbildung in Anlehnung an [8]

2.4.3 Forschungsansätze für Navigationssysteme

Die Personalisierung von Komponenten der Navigation ist ein relativ junges Forschungsfeld und ist erst mit der Verbreitung der Navigationssysteme aufgekommen. Arbeiten wie von Vollrath [95] passen die Zielführung auf Basis von Erkenntnissen der Verkehrspsychologie an. Die meisten Arbeiten konzentrieren sich allerdings auf die Routingkomponente. Es gibt zwei Schwerpunkte. Die Zielfunktion bildet den ersten Schwerpunkt. Hierbei wird analysiert, welche Ziele ein Fahrer beim Auswählen einer Route verfolgt. Der zweite Schwerpunkt befasst sich mit den Grundlagen, auf denen der Fahrer das Ziel verfolgt, analog zum Navigations-

system, das auf einem Graphen mit Gewichten eine Route mit minimaler Zielfunktion berechnet. Im Folgenden werden einige Ansätze präsentiert und in den Forschungskontext eingeordnet. Ebenso werden Nutz- und Anwendbarkeit für das zu entwickelnde Verfahren evaluiert.

Der Ansatz von Pang [73] beschäftigt sich mit der Forschungsfrage, wie man die Zielkriterien an den Fahrer anpassen kann. Dabei werden auf die Unzulänglichkeiten hingewiesen, die heute verfügbare Kriterien wie „kurz“ oder „schnell“ haben, wenn es darum geht, menschliche Entscheidungskriterien bei der Routenauswahl darzustellen. Dafür wird der Begriff der „Optimalität“ einer Route neu definiert und der Lösungsweg dieser optimalen Route skizziert. Dieser basiert, um den menschlichen nichtdeterministischen Entscheidungsweg besser modellieren zu können, auf *Fuzzy-Logik*.

Mit Fuzzy-Logik können vor allem Unsicherheiten modelliert sowie Unschärfen von umgangssprachlichen Beschreibungen dargestellt werden. Sie ist eine Verallgemeinerung der *zweiwertigen Booleschen Logik*. Beispielsweise kann damit die sogenannte Fuzziness von Angaben wie „ein bisschen“ „ziemlich“ oder „stark“ mathematisch in Modellen erfasst werden und eignet sich daher hervorragend für die Modellierung von menschlichen Entscheidungswegen [97]. Die Fuzzy-Logik basiert auf den Fuzzy-Mengen und sogenannten Zugehörigkeitsfunktionen, die Objekte auf Fuzzy-Mengen abbilden, sowie passenden logischen Operationen auf diesen Mengen. Bei technischen Anwendungen müssen außerdem Methoden zur Fuzzyfizierung und Defuzzyfizierung betrachtet werden, das heißt Methoden zur Umwandlung von Angaben und Zusammenhängen in Fuzzy-Logik und wieder zurück.

Fuzzy-Logik wird heute in unterschiedlichen Bereichen eingesetzt, z.B. in Bereichen der Automatisierungstechnik und der Betriebswirtschaft [73, 97]. Nützlich ist die Verwendung von Fuzzy-Logik oft dann, wenn keine mathematische Beschreibung eines Sachverhaltes oder Problems vorliegt sondern nur eine verbale Beschreibung. Auch wenn - wie fast immer - das vorhandene Wissen Lücken aufweist oder teilweise veraltet ist, bietet sich der Einsatz von Fuzzy-Logik an, um noch zu einer fundierten Aussage über einen aktuellen oder künftigen Systemzustand zu gelangen. Anschließend wird aus sprachlich formulierten Sätzen und Regeln mittels Fuzzy-Logik eine mathematische Beschreibung gewonnen, die in Rechnersystemen genutzt werden kann.

Interessant ist dabei, dass mit der Fuzzy-Logik auch dann Systeme sinnvoll gesteuert (bzw. geregelt) werden können, wenn ein mathematischer Zusammenhang zwischen den Ein- und Ausgabegrößen eines Systems nicht darstellbar ist oder

nur mit großem Aufwand erfolgen könnte, was bei einer menschlich gesteuerten Routenfindung durchaus der Fall sein kann. Daher wird in der Arbeit von Pang dieser Ansatz gewählt, um unscharfe Kriterien für die Auswahl einer Route zu finden. Der Ansatz hat den Vorteil, dass so den Kriterien für die Modellierung der Routenfindung mehr Spielraum eingeräumt wird.

Allerdings werden bei dem Ansatz die Diskrepanz zwischen dem Wissen des Navigationssystems und das des Fahrers nur bedingt berücksichtigt. Zudem werden durch die Fuzzifizierung und Defuzzifizierung der Entscheidungskriterien die Auswahlkriterien mehrmals verändert, so dass es zu einer Routenauswahl kommen kann, die der Fahrer nicht beabsichtigt, abgesehen davon, dass der Fahrer selbst meist eine subjektive Bewertung der Route vornimmt. Daher wirft dieser Ansatz durchaus interessante Fragen auf, lässt aber noch viel Spielraum für Interpretationen, so dass ein Einsatz in einem deterministischen System zur Bestimmung eines kürzesten Weges problematisch wäre. Vielmehr würde es sich anbieten, die Entscheidungen des menschlichen Fahrers entsprechend der verfügbaren Kriterien zu objektivieren, um den Fahrer eine bessere Bewertung der Routenauswahl zu ermöglichen.

Einen sehr intensiven und methodischen Einblick in das Routenauswahlverfahren gibt Li [56] in seiner Ausarbeitung. Dort werden anhand von Realdaten unterschiedliche Routen bewertet und Zielkriterien daraus abgeleitet. Es werden zuerst mögliche Einflusskriterien bestimmt und geordnet, siehe Abbildung 2.38.

Traveler	Age, gender, life cycle, income level, education, household structure, race, profession, length of residence, number of drivers in family, number of cars in family, etc.	
Route	Road	Travel time, travel cost, speed limits, waiting time. Type of road, width, length, number of lanes, angularity, intersections, bridges, slopes, etc.
	Traffic	Traffic density, congestion, number of turns, stop signs, and traffic lights, travel speed, probability of accident, reliability and variability in travel time, etc.
	Environment	Aesthetics, land use along route, scenery, easy pick-up/drop-off, safety, parking, etc.
Trip	Trip purpose, time budget, time of the trip, mode use, number of travelers	
Circumstances	Weather conditions, day/night, accident en route, route and traffic information, etc.	

Abbildung 2.38: *Routenauswahlkriterien, Abbildung aus [56]*

Darauf aufbauend werden die im Rahmen eines Projektes aufgezeichneten Routen untersucht, inwieweit sie Einfluss auf das Routenauswahlverfahren haben. Ebenso werden Fahrzeiten untersucht, um eine gefahrene Route zu bewerten. So werden sowohl die inneren Zielkriterien wie auch die äußeren Einflüsse in Relation gebracht und deren Relevanz anhand von Realdaten festgestellt.

Einen anderen Ansatz wählt McGinty [62] in ihrem Ansatz, einem Fahrer eine personalisierte Route anzubieten. Dort wird der Fahrer abhängig von seinem bisher gezeigten Verhalten in eine vorher festgelegte Klasse von Fahrern mit bestimmten Eigenschaften und Präferenzen eingeordnet. Diese Einordnung wird bei einer Routenberechnung genutzt, um eine den Eigenschaften der Klasse gerechten Route zu berechnen und vorzuschlagen. Dieses System ist an das Collaborative Filtering angelehnt. Hier sei anzumerken, dass dieses Verfahren der Einordnung in eine Klasse nicht unbedingt eine individuell optimale Lösung zur Folge haben kann. Zum anderen basiert das Konzept darauf, dass die erhobenen Daten kommuniziert und miteinander verglichen werden, was in der Praxis schwierig umzusetzen ist (siehe Abschnitt 2.4.1).

Anders geht Park [74] bei seinem Ansatz eines Entscheidungsbaum-Lernverfahren vor. Dort wird eine einmal getätigte Einordnung anhand des Routenauswahlverhaltens des Fahrers immer wieder überprüft und gegebenenfalls angepasst. Dieser iterative Ansatz ermöglicht eine fortwährende Verbesserung, was einerseits zu einer optimalen Lösung führt, andererseits aber auch für Verwirrung beim Fahrer und damit einer eventuell verringerten Akzeptanz führen kann. Daher ist die Bildung eines eindeutig nachvollziehbaren Modells, auf dem eine Personalisierung durchgeführt wird, einem solchen Ansatz vorzuziehen, der zwar eventuell schneller auf Fehlanpassungen reagieren kann, dafür aber eine zu große Diskontinuität aufweist.

In der Untersuchung von Yao [101] wird mobile Navigation, insbesondere Fußgänger-navigation betrachtet. Dabei werden Psychologie und die Untersuchung geographischer Daten miteinander in Beziehung gesetzt. Zudem werden die Fragestellungen untersucht, wie eine Anpassung erfolgen kann und vor allem was passiert, wenn diese ungenügend ist oder nicht den Wünschen des Nutzers entspricht. Dies wirft ein Licht auf die Frage nach der richtigen Einschätzung des Nutzers und seines gezeigten Verhaltens.

In den Arbeiten von Ehrenpfordt [25, 26] wird eine belastungsangepasste personalisierte Routenberechnung durchgeführt. Dabei wird zuerst das Kriterium der Belastung als ein weiteres Zielkriterium der Routenberechnung eingeführt und der Herzschlag als ein Maß dafür bestimmt. Darauf aufbauend wird ein Modell des

Fahrers erstellt, indem bestimmte Situationen und Klassen Belastungen zugeordnet werden. Somit ergeben sich neue Gewichte zur Routenberechnung, die nun Eingang finden können.

Die so berechneten Routen werden mit standardisierten Routen bezüglich der ursprünglichen Kriterien und der neuen Kriterien auf Basis von Feldversuchen verglichen. Ein Schema der neu gewählten Zielkriterien ist in Abbildung 2.39 gegeben. Dabei gibt es zwei Ausprägungen des Zielkriteriums Belastung. Zum einen wird versucht diese in der *entspannten Route* zu minimieren, zum anderen wird versucht, diese bei der *aktiven Route* in einem bestimmten Intervall zu halten und den Fahrer durch Auswahl der Route zu be- aber nicht zu überlasten.

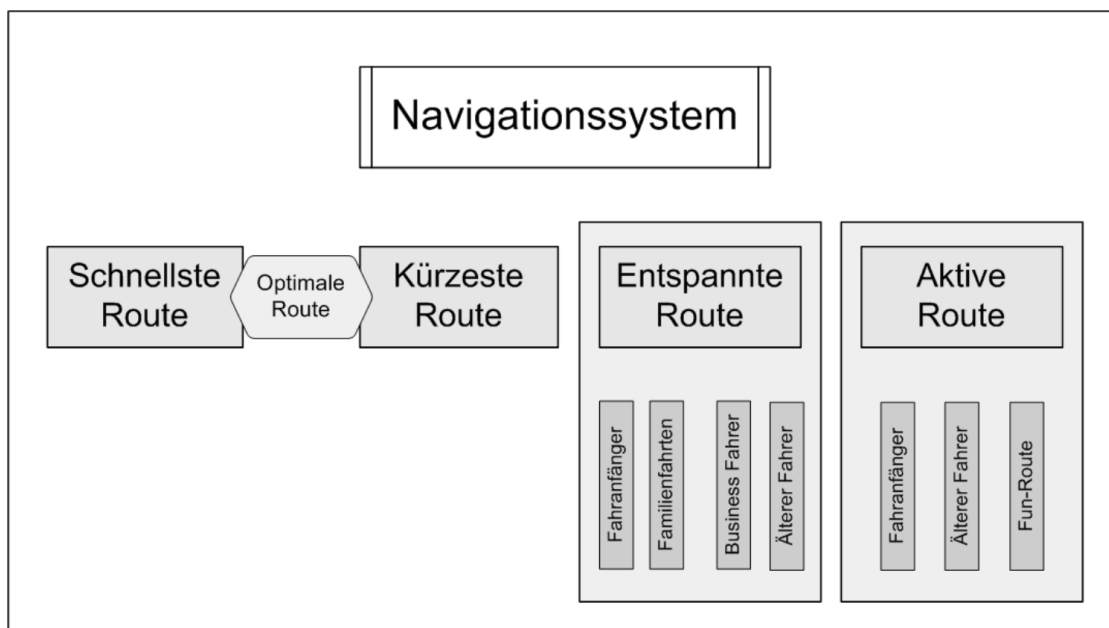


Abbildung 2.39: Einführung neuer Zielkriterien, Abbildung aus [26]

Die angewandte Methodik stellt einen sehr interessanten Ansatz dar, die Routenberechnung zu personalisieren, da hier für jeden Fahrer ein personalisiertes Belastungsprofil erstellt und für die Berechnung benutzt wird. Zudem wird das Verfahren implementiert und validiert. Allerdings wird dieses für ein neu eingeführtes Zielkriterium durchgeführt. Gewichtungen für klassische Kriterien für eine Personalisierung sind ungleich schwerer zu identifizieren. Außerdem muss darauf geachtet werden, diese Zielkriterien, die für die Fahrer von Bedeutung sind, durch die Zielfunktion der Belastung nicht zu sehr zu verschlechtern und so eine inakzeptable Route zu berechnen. Dieses würde nur auf geringe Akzeptanz seitens des Fahrers stoßen. Insbesondere dann, wenn zwar das Niveau der Belastung für die Route verringert, die Routendauer oder Routenlänge dafür deutlich erhöht wird.

2.4.4 Bewertung der Ansätze und Verfahren

Für die Umsetzung einer Personalisierung der Routingkomponente von Navigationssystemen wurden verschiedene modellgestützte Ansätze und Methoden aus unterschiedlichen Bereichen untersucht. Ausgangspunkt waren im Navigationssystem implementierte Systeme mit adaptierbaren Funktionen bzw. adaptiven Ansätzen basierend auf einer rudimentären Modellierung.

Am weitesten verbreitet und entwickelt sind Personalisierungsansätze im IT-Bereich. Ansätze im E-Commerce ordnen verschiedene Nutzer verschiedenen Gruppen mit ähnlichen Eigenschaften zu. Hierfür sind eine Vielzahl von Nutzern notwendig. Diese Anzahl zu erreichen ist im Fahrzeug- bzw. Navigationsbereich aufgrund der Sensibilität der Daten nicht praktikabel. Ob Ansätze aus dem Verkehrsbereich (siehe Abschnitt 2.3.2) genutzt werden können, ist fraglich, da für eine Personalisierung die Anonymisierung nicht aufrechterhalten werden kann. Vielmehr sind Systeme zu nutzen, die nur den Fahrer selbst analysieren und nicht in Zusammenhang mit anderen Fahrern sondern mit vorab definierten Eigenschaften stehen.

Im Bereich des E-Learnings wird dieser Ansatz verfolgt. Die Leistungen des Nutzers werden analysiert und entsprechend angepasst. Allerdings muss hierfür nur der Nutzer und seine Interaktion mit dem System berücksichtigt werden. Umwelteinflüsse, die auf das Verhalten im Straßenverkehr einen großen Einfluss haben, können nicht abgebildet werden. Somit können die weit entwickelten Ansätze aus dem IT-Bereich nicht auf die Situation im Fahrzeug übertragen werden.

Den Fahrer und seine Eigenschaften zu modellieren und für eine Personalisierung von Systemen im Fahrzeug zu nutzen, wird zunehmend im Bereich der FAS, z.B. bei Bremsassistenten, durchgeführt. Allerdings analysieren diese Modelle das Verhalten auf den unteren Ebenen der Fahraufgabe und berücksichtigen nicht das Verhalten auf der Navigationsebene. Zwar gibt es Zusammenhänge zwischen diesen Ebenen, doch zur Untersuchung des Routenauswahlverhaltens ist es notwendig, ein Modell auf der Ebene der Navigation zu erstellen.

Aufgrund der geringen Übertragbarkeit existierender Systeme zur Personalisierung von Navigationssystemen wurden Forschungsansätze zur Personalisierung untersucht. Es hat sich gezeigt, dass bereits einige Ansätze existieren, diese aber für sich stehen und nur gewisse Teilbereiche abdecken. Zudem offenbaren sie Unzulänglichkeiten, die die Akzeptanz seitens des Fahrers deutlich schmälern könnten. Auch werden oftmals Alternativzielkriterien bzw. -routingverfahren untersucht, deren Anwendbarkeit auf heutzutage verfügbare Navigationssysteme nicht gesichert ist.

In dieser Arbeit werden daher die evaluierten Ansätze nicht für die Modellierung oder die Personalisierung genutzt bzw. darauf aufgebaut. Vielmehr werden die Arbeiten als Anregungen genutzt, um ein eigenständiges, umfassendes Verfahren zur Personalisierung des bestehenden Routingverfahrens auf Basis eines Modells des Fahrers zu entwickeln und zu implementieren. Es werden die Entscheidungsgrundlagen sowie -prozesse der Routenauswahl analysiert und ein Modell der navigatorischen Präferenzen eines Fahrers erstellt. Diese bilden die Grundlage für eine Personalisierung der Routingkomponente von Navigationssystemen.

3 Theoretischer Ansatz für eine modellgestützte Personalisierung

In diesem Kapitel wird der Ansatz für eine modellgestützte Personalisierung der Routingkomponente von Navigationssystemen sowie das zugrundeliegende Konzept vorgestellt. Das Ziel der Personalisierung ist es, die existierende Routingkomponente von Navigationssystemen zu verbessern. Für die Entwicklung sowie die Umsetzung des dazugehörigen Verfahrens stellen sich die folgenden Forschungsfragen:

- Welche Daten über den Fahrer werden für die Analyse navigatorischer Präferenzen des Fahrers benötigt?
- Mit welchen Methoden werden Informationen aus den Daten extrahiert?
- Wie kann Wissen generiert und darauf aufbauend ein Modell der navigatorischen Präferenzen des Fahrers erstellt werden?
- Wie erfolgt die Personalisierung der Routingkomponente von Navigationssystemen auf Basis des Modells?
 - Ist das Modell quantifizierbar und eine Anwendung auf die Routingkomponente möglich?
 - Welche Ziele werden mit der Personalisierung verfolgt?
 - Wie sieht ein entsprechendes Verfahren aus?
- Wird eine Verbesserung durch die Personalisierung erreicht, und wie kann sie gemessen werden?

Im Folgenden werden die Vorgehensweise und die zugrundeliegenden Ideen zur Beantwortung der Forschungsfragen skizziert. Eine detaillierte Beschreibung der einzelnen Schritte des Verfahrens sowie der dazugehörigen Ergebnisse erfolgt in den anschließenden Kapiteln.

3.1 Personalisierung der Routingkomponente

3.1.1 Entwicklung des Konzepts

Wie in Kapitel 2 dargestellt, betreffen die relevanten einstellbaren Zielkriterien der Routingkomponente von Navigationssystemen die Routendauer sowie die Routenlänge [80]. Zusätzlich gibt es die Möglichkeit, beide Zielkriterien miteinander zu kombinieren. Diese Zielkriterien liegen im Allgemeinen auch dem Routenauswahlverhalten des Fahrers zugrunde. Allerdings kann es bei der Gewichtung der beiden Kriterien zu Unterschieden zwischen einem Navigationssystem und dem Fahrer sowie zwischen den Fahrern kommen [11].

So ist es sehr unwahrscheinlich, dass ein menschlicher Fahrer eine Route mit einem Umweg von mehreren Kilometern auswählt, um wenige Sekunden einzusparen. Ist dem Fahrer die Route bekannt, so wird er vermutlich die zeitlich längere, aber streckenmäßig deutlich kürzere Route wählen. Auch bei einer unbekannten Strecke würden Fahrer eine solche Route, die mit Hilfe einer Straßenkarte ausgewählt wird, einer zeitlich nur unwesentlich kürzeren und dafür streckenmäßig deutlich längeren Route vorziehen [11]. Eine Routenberechnung, die ausschließlich die Zeit als Zielkriterium berücksichtigt, würde die streckenmäßig deutlich längere Route und so eine für den menschlichen Fahrer inakzeptable Route auswählen.

Dieses Beispiel zeigt anhand der Routendauer sowie der Routenlänge Unterschiede zwischen dem System und dem Fahrer und den potentiellen Nutzen einer Personalisierung auf (siehe Abbildung 2.27). Die Zielkriterien können unterschiedlichen Dimensionen zugeordnet werden. Die Routendauer gehört zu der Dimension der Zeit, die Routenlänge zu der Dimension des Raumes.

Dabei stellt sich vor allem die Frage nach der Substitution zwischen Routenlänge und -dauer. Stellt das Beispiel noch einen anschaulichen Extremfall dar, so ist diese Frage bei ähnlichen Ersparnissen bzw. Mehraufwänden nicht eindeutig zu beantworten. Hier kann es zu Unterschieden zwischen den Fahrern in Abhängigkeit ihrer Präferenzen kommen. Während einige Fahrer großen Wert auf eine kurze Routendauer legen, so gibt es auch Fahrer mit einer Präferenz für eine kurze Routenlänge.

Einige Navigationssysteme bieten die Möglichkeit an, manuell Gewichte für Routenlänge und -dauer und somit für die Dimensionen der Zeit und des Raumes zu setzen, wobei es zu den in Abschnitt 2.3.4 dargestellten Problemen durch Fehleinschätzungen des Fahrers kommen und zu einer im Sinne des Fahrers unerwünschten Route führen kann. Eine Adaption des Systems erfolgt bisher aus-

schließlich für die Dimension der Zeit über eine Analyse der gemessenen Geschwindigkeiten (siehe Abschnitt 2.3.5). Daraus folgt eine automatische Anpassung der angenommenen Geschwindigkeit für die verschiedenen Kanten und somit eine Veränderung der berechneten Routen.

Dabei wird der Wert, den der Fahrer auf die Routenlänge und damit der Dimension des Raumes legt, nicht berücksichtigt. Somit geht auch eine mögliche Substitution zwischen den beiden Dimensionen nicht in die Berechnung mit ein, obwohl dieses ein wichtiges Entscheidungskriterium des Fahrers bei der Routenauswahl ist.

Somit muss es Ziel sein, das Verhältnis zwischen Raum und Zeit für jeden Fahrer abzubilden und der Routingkomponente eines Navigationssystems zur Verfügung zu stellen. In Abbildung 3.1 ist dieses Konzept als *verschiebbarer Regler* illustriert. Dieser gibt die Gewichtung zwischen den Dimensionen Raum und Zeit an. Für jeden Fahrer werden beide Dimensionen personalisiert mit einem Anteil gewichtet. Der Raumanteil kann dabei als fixer Bestandteil angesehen werden, da sich Längen nicht verändern können. Der Zeitanteil hingegen ist variabel. Die Variabilität ergibt sich aus den unterschiedlichen gemessenen Geschwindigkeiten der Fahrer und zeigt die Präferenz des Fahrers für die Dimension der Zeit an.

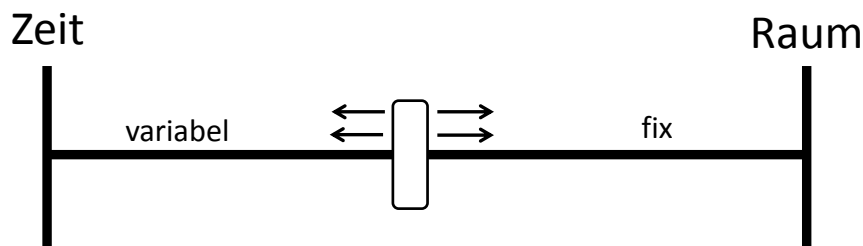


Abbildung 3.1: Veranschaulichung der Gewichtung der Dimensionen Raum (*fix*) und Zeit (*variabel*) bei der Routenberechnung

Wird durch Regelung der Fokus auf die Dimension des Raumes gelegt, so entspricht dies einer kürzesten Route. In diesem Fall ist die Routenberechnung *fix*, und es ergibt sich für alle Fahrer die gleiche (kürzeste) Route. Im umgekehrten Fall hingegen liegt der Fokus auf der Zeit und entspricht einer schnellsten Route. In diesem Fall kann für jeden Fahrer eine Berechnung auf der Basis der personalisierten Kantengewichte und damit *variabel* erfolgen. Entsprechend wird nur die Dimension der Zeit berücksichtigt.

Wird für jeden Fahrer der Regler personalisiert eingestellt, so wird auch die Zielfunktion der Routenberechnung entsprechend personalisiert. Auf diese Weise wird nicht nur die Dauer (Dimension Zeit) sondern auch die Länge (Dimension Raum)

mit den entsprechenden Gewichten in die Berechnung einbezogen, mit dem Ziel, eine Route zu berechnen, die die Ziele des Fahrers in größerem Maß erreicht.

3.1.2 Methoden der Bewertung

Für die Quantifizierung einer potentiellen Verbesserung durch die Einbeziehung der Dimension der Zeit ist ein Gütemaß zu definieren. Generell sollte eine Personalisierung zur Folge haben, dass für eine Start-Zielkombination die vom Navigationssystem berechnete personalisierte Route (*Individualroute*) mit der vom Fahrer ausgewählten Fahrer-Realroute übereinstimmt. Ausgehend von diesem absoluten Maß ist somit eine Verbesserung erreicht, wenn die Übereinstimmung der Individualroute mit der Fahrer-Realroute größer ist als die Übereinstimmung der vom Navigationssystem berechneten Standardroute mit der Fahrer-Realroute. Als Standardroute wird die schnellste Route verwendet. In Abbildung 3.2 ist das Gütemaß der Übereinstimmung auf Basis der Kanten illustriert. Jede Route besteht aus einer Menge an Kanten. Für die Bestimmung der Güte werden die Schnittmengen zwischen der Fahrer-Realroute und den beiden berechneten Routen genutzt. Bezogen auf die Gesamtkantenmenge der Fahrer-Realroute kann so der Anteil übereinstimmender Kanten bestimmt werden.

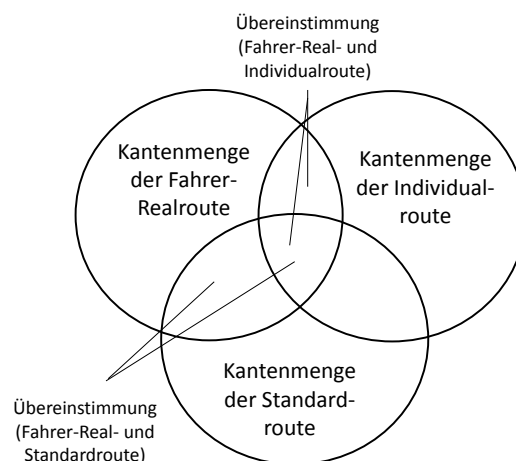


Abbildung 3.2: *Bewertung auf Basis der Übereinstimmung*

Für ein weiteres Gütemaß werden die Streckencharakteristika einer Route analysiert. Zur Charakterisierung der Strecke werden die Anteile der verschiedenen Straßenklassen genutzt. Eine Strecke unterteilt sich in die drei Straßenklassen innerorts, Überland sowie Autobahn. Ziel ist es, dass die Streckencharakteristika der

Individualroute ähnlich zu denen der Fahrer-Realroute sind. So sollte z.B. einem Fahrer, der generell wenig auf der Autobahn fährt, auch eine Route angeboten werden, die keine oder nur wenig Autobahnabschnitte enthält. Dieser Ansatz ist an die manuelle Einstellungsmöglichkeit von Autobahn (erlaubt, vermeiden, verboten) aus Abschnitt 2.3.4 angelehnt. Auch für das Gütemaß der Streckencharakteristika ist eine Verbesserung erreicht, wenn die Anteile der einzelnen Straßenklassen der Individualroute mehr mit denen der Fahrer-Realroute übereinstimmen als es für die Standardroute der Fall ist.

In Abbildung 3.3 ist der Ansatz illustriert. Sowohl für die Fahrer-Realroute wie auch für die berechnete Route (Standard oder Individual) werden die Längenanteile der drei Straßenklassen an der Gesamtlänge bestimmt. Die Anteile werden miteinander verglichen und aggregiert, um ein Gütemaß dafür zu erhalten wie ähnlich sich die Routen bezüglich der Streckencharakteristika sind.

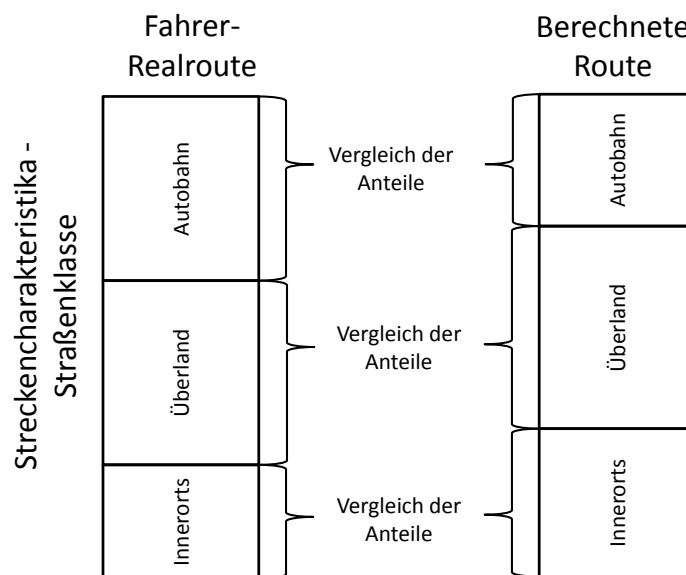


Abbildung 3.3: *Bewertung auf Basis der Straßenklasse*

Wie bereits dargestellt, wählt der Fahrer ebenso wie das Navigationssystem eine Route unter Berücksichtigung eines Zielkriteriums aus. Dies ist zum Beispiel die Minimierung der benötigten Zeit (schnellste Route). Durch eine Personalisierung sollte die Erreichung der Ziele verbessert werden.

Daher wird als drittes und wichtigstes Maß der Vergleich der Routencharakteristika bezüglich des bzw. der Zielkriterien genutzt. Mit Hilfe dieses Maßes der Zielkriterien lässt sich eine Verbesserung durch die Personalisierung im Vergleich zur Standardroute direkt quantifizieren und so der Nutzen der Personalisierung bestimmen.

In Abbildung 3.4 ist der Ansatz der Bewertung für das Gütemaß der Zielkriterien dargestellt. Sowohl Standard- wie auch Individualroute erreichen die vom Fahrer verfolgten Ziele in einem bestimmten Ausmaß. Die Zielkriterienerreichung der beiden berechneten Routen wird auf Basis der gleichen Datenbasis quantifiziert und miteinander verglichen. Ziel ist es, die Zielfunktionswerte für die personalisierten Routen gegenüber der Standardroute zu verbessern.

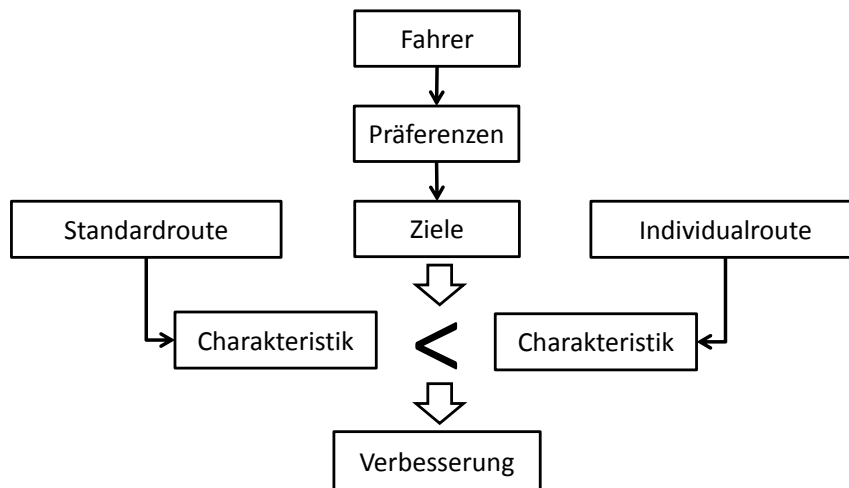


Abbildung 3.4: *Bewertung auf Basis der Zielkriterien*

3.2 Modellierung der navigatorischen Präferenzen

3.2.1 Anforderungen

Als Grundlage für die Personalisierung wird ein Modell erstellt, das auf den Dimensionen der Zeit sowie der des Raumes basiert. Ebenso von Bedeutung ist die Anwendbarkeit des Modells auf die Routenberechnung von Navigationssystemen. Eine narrative Beschreibung der Eigenschaften eines Fahrers, wie z.B. in [56] ist zwar hilfreich zur Einordnung des Fahrers, doch ist die Nutzbarkeit für eine Routenberechnung gering. Wie in Abschnitt 2.2.4 beschrieben, basieren Routenberechnungsverfahren von heutigen Navigationssystemen auf einem KW-Verfahren, das auf einem Graphen mit Kantengewichten durchgeführt wird. Die Kantengewichte des Graphen können dabei durch die Länge aber auch durch die Zeit, die für die Durchfahung der Kanten benötigt wird, definiert sein. Alternative Routingverfahren, wie bei Pang und Park [73, 74] beschrieben, haben sich nicht durchgesetzt. Aus diesem Grund muss eine Modellierung navigatorische Präferenzen quantifizie-

ren und in eine Neubewertung der Kanten bzw. in die Erstellung einer Zielfunktion einfließen lassen. Nur in diesem Fall ist das Modell für die Routenberechnung heutiger Navigationssysteme nutzbar.

Grundlage der Modellierung sind die Analysen des navigatorischen Verhaltens der Fahrer. Wie bereits im Abschnitt 2.4.2 gezeigt, lag der Schwerpunkt für Fahrermodelle bisher auf den unteren Ebenen der Bahnführung und Stabilisierung. Dementsprechend gibt es bisher nur wenige Modelle, auf die sich eine Modellierung der navigatorischen Präferenzen eines Fahrers stützen kann. Dazu ist es notwendig, eine solche Modellierung auf der Ebene der Navigation empirisch mit Hilfe geeigneter Daten, die die navigatorischen Präferenzen beschreiben, zu belegen. Ein entsprechender Ansatz sollte mehrere Aspekte der Navigation berücksichtigen und diese in ein fundiertes Modell einarbeiten. Somit ergeben sich vier entscheidende Anforderungen an die Modellierung, die in Abbildung 3.5 nochmals grafisch dargestellt sind.

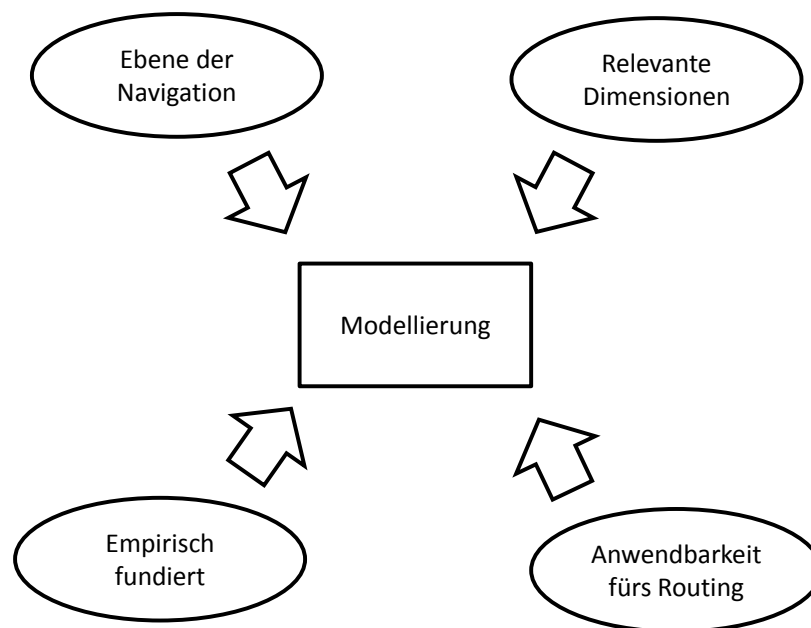


Abbildung 3.5: Anforderungen an die Modellierung

3.2.2 Dimensionen und Perspektiven der Modellierung

Es wird ein zweidimensionales Modell der navigatorischen Präferenzen eines Fahrers erstellt, das sowohl die Dimension der Zeit wie auch die Dimension des Raumes berücksichtigt. Die Modellierung basiert auf drei Analyseperspektiven. Die erste Perspektive umfasst die Kanten und ihre Gewichte (Kantenperspektive). Die

Routen werden für die zweite Perspektive (Routenperspektive) analysiert. Für die dritte Perspektive werden die Gesamtheit aller Routen, die in einem Zeitraum gemessen wurden und ein Netzwerk bilden, einer Analyse unterzogen und bewertet (Netzwerkperspektive).

Für die einzelnen Perspektiven werden verschiedene Maße untersucht, die den Dimensionen der Zeit und des Raumes zugeordnet werden können. Für die Kantenperspektive wird das *Maß der Geschwindigkeit* M_v (Modell 1) untersucht, das zur Dimension der Zeit gehört. Für die Routenperspektive werden die Maße *Dauer* M_d (Dimension Zeit) und *Länge* M_l (Dimension Raum) analysiert und entsprechend zwei Modelle (Modell 2a und 2b) erstellt. Das Maß der *Abdeckung* M_a ergibt sich aus Analysen für die Netzwerkperspektive (Modell 3). Eine Übersicht über analysierte Perspektiven und Dimensionen bzw. Maße ist in Tabelle 3.1 dargestellt.

Dimension	Perspektive	Kanten	Routen	Netzwerk
	Maß			
Zeit	Geschwindigkeit	Modell 1		
	Dauer		Modell 2a	
Raum	Länge		Modell 2b	
	Abdeckung			Model 3

Tabelle 3.1: *Perspektiven und Dimensionen der Modellierung*

Für die Erstellung eines zweidimensionalen Modells werden im nächsten Schritt die Beziehungen der Maße zueinander untersucht. Eine Visualisierung der Zusammenhänge findet sich in Abbildung 3.6. Die Maße Geschwindigkeit und Dauer gehören zur Dimension der Zeit (linke Seite) und die Maße Länge und Abdeckung zur Dimension des Raumes (rechte Seite). Dabei wurden die Dauer und die Länge für die gleiche Perspektive (Routenperspektive, oben) erstellt.

Die Maße, die die gleiche Dimension betreffen, sollten jeweils in Zusammenhang zueinander stehen und eine Korrelation aufweisen. So sollte z.B. ein Fahrer, der schnell fährt, auch kürzere Routendauern erreichen. Gleiches sollte für Maße, die für die gleiche Perspektive erstellt wurden, gelten. Für die Maße, die für verschiedene Perspektiven ermittelt worden sind und unterschiedlichen Dimensionen zugeordnet werden, ist der Zusammenhang und eine entsprechende Korrelation weniger naheliegend. Hier wird analysiert, ob und unter welchen Umständen eine Korrelation gefunden werden kann. Darüber hinaus wird auch die Frage untersucht, inwieweit dies für eine Modellierung genutzt werden kann.

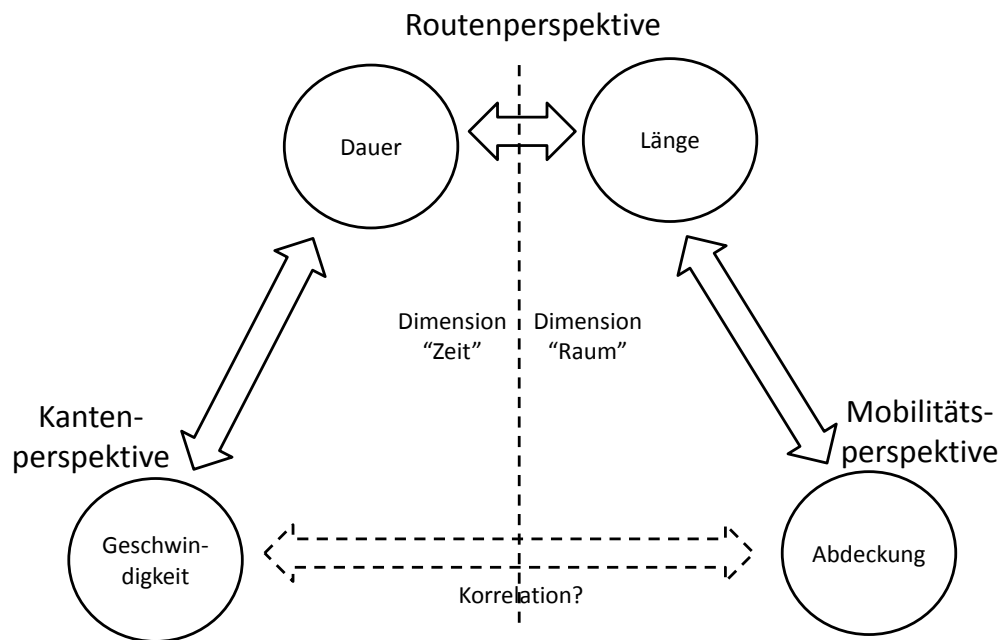


Abbildung 3.6: Visualisierung der potentiellen Dimensions- und Perspektivenkorrelationen

Das anschließende Ziel ist es, auf Basis der Maße der verschiedenen Perspektiven ein zweidimensionales Modell des Fahrers zu erstellen, das beide Dimensionen berücksichtigt. Eine deskriptive Beschreibung ist in Abbildung 3.7 dargestellt. Generell können sich vier verschiedene *Routingtypen* ergeben, wenn die beiden Dimensionen unabhängig voneinander betrachtet werden. Fahrer mit hoher Präferenz sowohl für die Zeit wie auch für den Raum werden als *Geschäftsfahrer* klassifiziert.

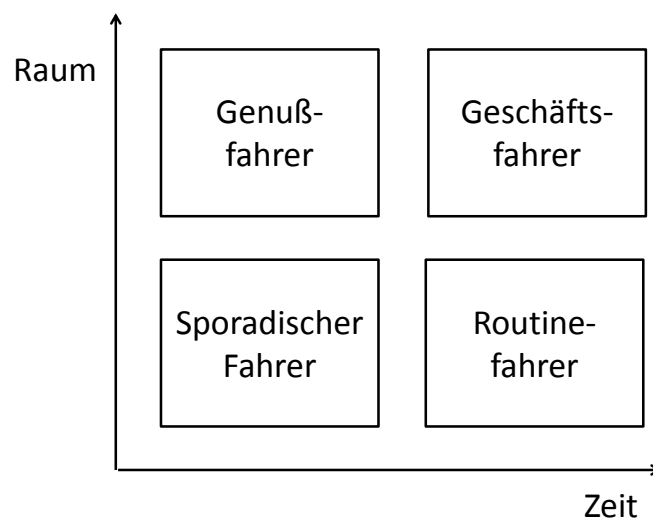


Abbildung 3.7: Deskriptive Beschreibung der Routingtypen

Routinefahrer haben zwar ebenfalls eine hohe Präferenz für die Dimension der Zeit jedoch im geringeren Maße für den Raum. Dahingegen haben *Genussfahrer* eine hohe Präferenz für die Dimension des Raumes aber weniger für die Dimension der Zeit. Fahrer, die für beide Dimensionen eine geringe Präferenz zeigen, werden als *sporadische Fahrer* klassifiziert.

3.3 Vorgehen in Anlehnung an einen Knowledge Discovery Process

Das für die Personalisierung entwickelte Vorgehen lehnt sich an einen KDP an. Ziel eines KDPs ist es, Wissen aus zumeist großen Datensätzen zu generieren. Ein wichtiger Schritt ist dabei das *Data Mining*, das häufig auch als Synonym für ein KDP genutzt wird. Allerdings sind dem Data Mining in einem KDP weitere Schritte vor- sowie nachgeschaltet. Mitte der neunziger Jahre wurde ein KDP erstmals von Fayyad et al. beschrieben [27]. Er besteht aus neun Schritten:

1. Entwicklung und Durchdringung des Anwendungsbereichs
2. Erstellung eines Zieldatensatzes
3. Datenbereinigung und -aufbereitung
4. Datenreduktion und -transformation
5. Definition der Data Mining Aufgabe
6. Auswahl des Data Mining Algorithmus
7. Durchführung des Data Minings
8. Interpretation der gefundenen Muster
9. Evaluierung und Einbettung des generierten Wissens

Eine Visualisierung der Datenverarbeitungsschritte ist in Abbildung 3.8 dargestellt. Ausgehend vom Anwendungsbereich werden die Daten selektiert und vorverarbeitet. Diese dienen als Grundlage für einen ausgewählten Data Mining Ansatz zur Mustererkennung. Die gefundenen Muster werden interpretiert, evaluiert und in Form von Wissen dem Anwendungsbereich zur Verfügung gestellt.

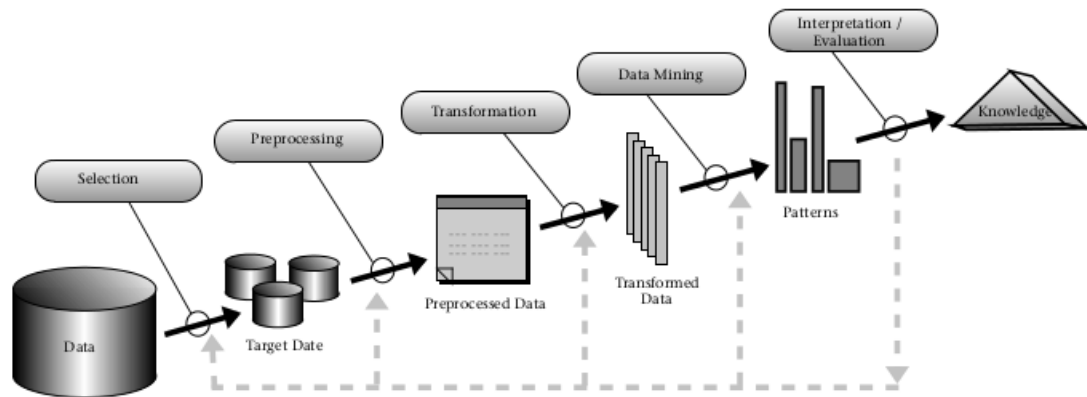


Abbildung 3.8: Überblick über die Datenverarbeitungsschritte eines KDPs, Abbildung aus [27]

Der Prozess ist iterativ angelegt, so dass es zu Mehrfachdurchläufen bzw. Wiederholung einzelner Schritte kommen kann. Darüber hinaus gibt es noch weitere Ansätze, wie das acht Schritte umfassende Modell von Anand und Buchner [2]. Diese Modelle werden den *wissenschaftlichen Ansätzen* zugeordnet, deren Hauptaugenmerk auf dem Generieren von Wissen liegt.

Darauf aufbauende *industrielle Ansätze* haben dazu das Ziel, das generierte Wissen für Applikationen bereitzustellen. Eines der am weitesten verbreiteten Ansätze ist der *CRoss-Industry Standard for Data Mining (CRISP-DM)* [59]. Das Vorgehen verläuft ähnlich wie wissenschaftliche Ansätze, allerdings wird der Schritt zur Generierung des Wissens, das Data Mining, allgemeiner als Modellierung angesehen. In Abbildung 3.9 ist das Vorgehen dargestellt. Es umfasst sechs Schritte:

1. Geschäftsverständnis
2. Datenverständnis
3. Datenaufbereitung
4. Modellierung
5. Evaluierung
6. Bereitstellung

Ausgehend vom *Business Understanding* sowie dem *Data Understanding* werden die Daten aufbereitet und dienen als Grundlage für die Modellierung. Dabei gibt es zwischen den meisten Schritten die Möglichkeit der Iteration. Die Modellie-

rung wird evaluiert, woraufhin die Ergebnisse bereitgestellt und integriert werden können. Alternativ kann der Prozess iterativ neu gestartet werden.

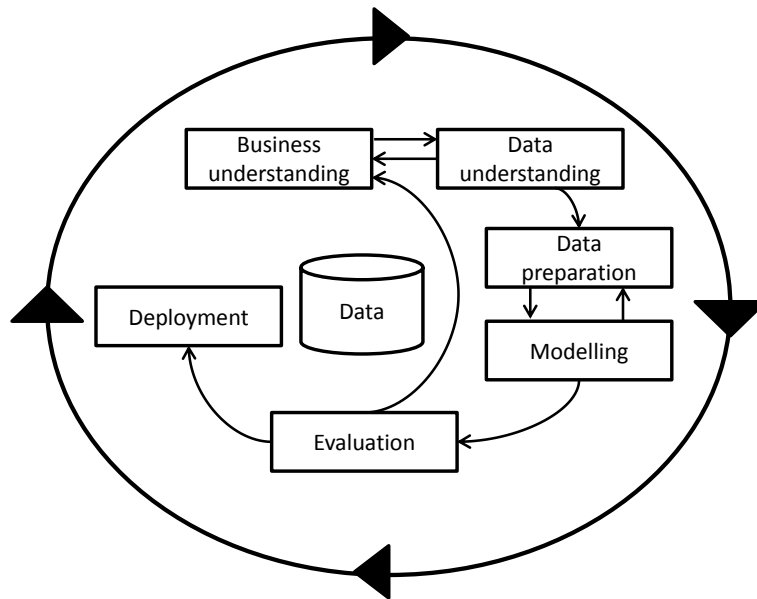


Abbildung 3.9: Aufbau des CRISP-DM KDPs, Abbildung nach [59]

Ein weiterer Ansatz ist das *Fünf-Schritte-Modell* von Cabena et al. [15]. Dazu gibt es *hybride Ansätze* [18], die Schritte der wissenschaftlichen und industriellen Ansätze miteinander kombinieren. Die meisten Ansätze sind ähnlich aufgebaut und umfassen das Verständnis für den Anwendungsbereich, das Data Mining sowie die Evaluierung des generierten Wissens.

Der in dieser Arbeit verwendete Ansatz zur Modellierung und Personalisierung mit seinen Schritten orientiert sich vornehmlich am CRISP-DM KDP. Die Darstellung der Funktionsweise von Navigationssystemen sowie das Ziel der Personalisierung zur Verbesserung entsprechen dem Geschäftsverständnis. Anforderungen an die benötigten Daten werden formuliert und auf die in dieser Arbeit genutzten Daten angewendet (Datenverständnis). Die Daten werden aufbereitet und dienen als Grundlage für die Modellierung. Die Evaluierung erfolgt durch die Durchführung und Bewertung der Personalisierung und ermöglicht zugleich die Nutzung für die Routingkomponente von Navigationssystemen (Bereitstellung).

Analog zum KDP läuft das Vorgehen in dieser Arbeit, siehe Abbildung 3.10. Grundlage bildet eine erhobene Datenbasis die das Routenauswahlverhalten von Fahrern beinhaltet. In einem ersten Schritt wird diese Datenbasis aufbereitet. Darauf baut die Modellierung inklusive der Analyse der drei Perspektiven auf. Dieses Modell

wird für die Personalisierung der Routingkomponente genutzt und bewertet, um den Nutzen der Personalisierung zu quantifizieren.

Eine detaillierte Beschreibung des gesamten Verfahrens ist in Kapitel 5 und die dazugehörigen Ergebnisse in Kapitel 7 zu finden. Die Aufbereitung der Datenbasis, die in Kapitel 4 vorgestellt und eingeordnet wird, ist detailliert in Kapitel 6 beschrieben.

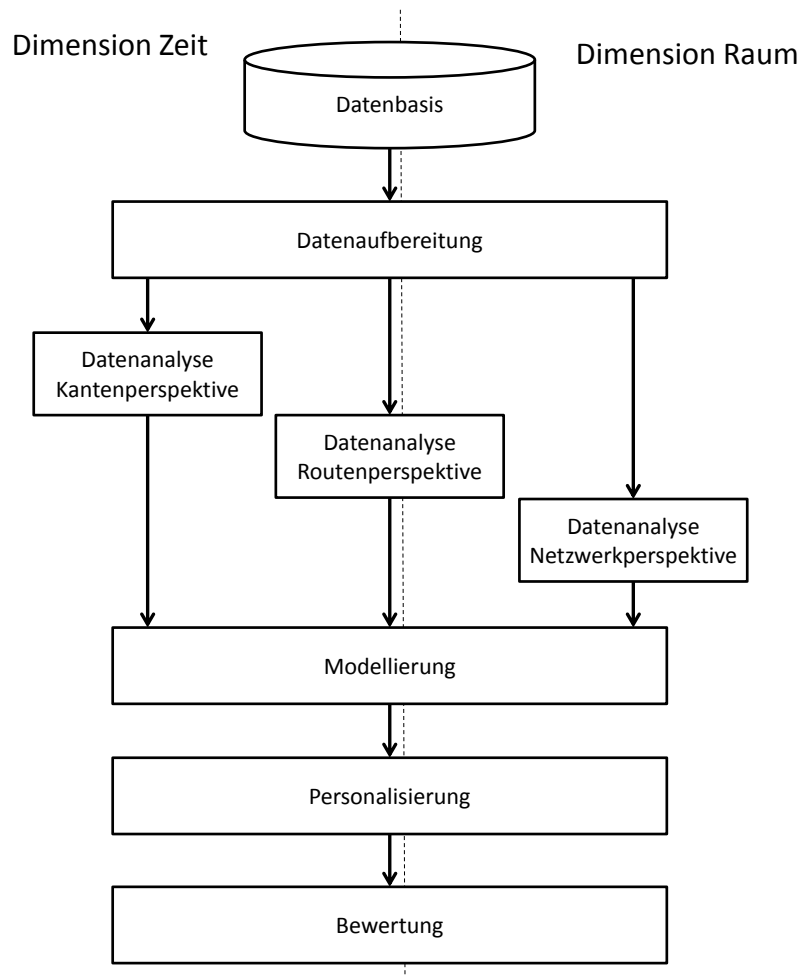


Abbildung 3.10: *Struktur des Vorgehens für die Modellierung und Personalisierung der navigatorischen Präferenzen eines Fahrers*

4 Datenbasis der empirischen Untersuchungen

Grundlage der Untersuchungen und des Verfahrens bilden Daten aus dem Projekt euroFOT [47, 63]. Ziel des europaweiten euroFOT-Projekts war es, Auswirkungen verschiedener Fahrzeugsysteme auf das Fahrerverhalten zu untersuchen. Während des Projektzeitraums 2010 bis 2012 wurden Daten in verschiedenen europäischen Ländern erhoben. Die folgende Darstellung des Projekts, sowie die Stichprobenbeschreibung und -einordnung wurde zum Teil aus Metz et al. [64] entnommen.

4.1 Beschreibung der Datenbasis

Für ein Teilprojekt mit dem Schwerpunkt „Navigationssysteme“ wurde 115 Fahrern ein Fahrzeug für einen Zeitraum von drei Monaten zur Verfügung gestellt. In einem Monat des Untersuchungszeitraums war es nicht erlaubt, ein Navigationssystem zu nutzen, während eingebaute bzw. mobile Navigationssysteme in den übrigen zwei Monaten zur Verfügung standen.

Damit stellt die Versuchsanordnung des Projekts ein *Field Operational Test (FOT)* dar. „FOTs gehören, wie auch *Naturalistic Driving Studies (NDS)*, unter methodischen Kriterien zur Gruppe der Fremdbeobachtungsmethoden mit stark bis vollständiger apparativer Datenerfassung. Bei einem FOT soll eine neue Funktion (oder mehrere) in ihrer Verwendung im Feld unter normalen Betriebsbedingungen überprüft werden. Üblicherweise wird ein quasi-experimentelles Design im Sinne von „Fahren mit vs. ohne Funktion“ verwendet. An Messgrößen wird eingeführt, was zu einer Beschreibung der untersuchten Funktion und ihres Gebrauchs notwendig ist. Aus dieser methodischen Struktur ergeben sich auch die Unterschiede zu einer NDS. Bei einer NDS gibt es während des Versuchszeitraums keine Vorgaben an den Fahrer, um das natürliche Fahrverhalten des Fahrers untersuchen zu können. Im FOT wird der Fahrer zumindest darüber instruiert, wann er die Funk-

tion nutzen soll und wann nicht. Der Fahrer ist sich deshalb deutlicher bewusst, unter Beobachtung zu stehen als dies bei einer NDS der Fall ist.“[64]

„Aus der Notwendigkeit, dass der Teilnehmer an einem FOT in die Untersuchungsfrage eingeweiht werden muss, ergibt sich eine Ausrichtung der Aufmerksamkeit des Fahrers auf die untersuchte Funktion und damit ein geringerer Grad an Naivität oder Unbefangenheit. Dies gilt in deutlich reduziertem Maße für alle anderen Fahr- und Bediencharakteristika, deren Grad an Reaktivität nicht anders ist als bei einer NDS.“[64] Im Projekt konnte der Fahrer Ziele sowie die entsprechende Route frei wählen. Die Routenberechnung durch das Navigationssystem konnte ebenfalls, bei Verfügbarkeit, frei genutzt werden. Somit eignen sich die Daten zur Untersuchung des Routenauswahlverhaltens der Fahrer.

4.1.1 Datenaufzeichnung

Es wurden zwei verschiedene Fahrzeugtypen verwendet. Es standen jeweils 15 Fahrzeuge für die Durchführung mehrerer Aufnahmewellen bereit. Die Datenaufzeichnung in den Versuchsfahrzeugen startete mit Öffnen der Fahrertür und endete mit Abstellen des Motors inklusive eines gewissen Nachlaufs. Sie beinhaltet die Aufzeichnung von Zeitverlaufsdaten aus den CAN- und anderen Fahrzeug-Bussen sowie von vier für den FOT verbauten Videokameras, siehe Tabelle 4.1.

	Fahrzeugtyp 1	Fahrzeugtyp 2
Anzahl Fahrzeuge	15	15
Anzahl Fahrer	60	51
Datenaufzeichnung: Fahrzeugdaten	4x CAN-Bus (10 Hz)	4x CAN-Bus (20 Hz) MOST (20 Hz) FlexRay (20 Hz)
Video	4 Kameras (5 Hz)	4 Kameras (5 Hz)

Tabelle 4.1: Übersicht über die Datenaufzeichnung, Tabelle nach [64]

Ein Visualisierung der Videobilder für beide Fahrzeugtypen ist in Abbildung 4.1 dargestellt. Zu Beginn einer Fahrt mussten die Fahrer durch Drücken einer Taste angeben, ob die entsprechende Route bekannt oder unbekannt ist. In den Fahrzeugen des Typs 1 war zusätzlich eine Taste installiert, mit der eine Audioaufzeichnung vom Fahrer gestartet werden konnte. Dies ermöglichte es den Fahrern, Kommentare zum aktuellen Fahrgeschehen abzugeben.



(a) Fahrzeugtyp 1

(b) Fahrzeugtyp 2

Abbildung 4.1: *Standbild der Videoaufzeichnung im Fahrzeugtyp 1 (links) und Fahrzeugtyp 2 (rechts), Darstellung aus [64]*

Die aufgezeichneten Verlaufsdaten enthalten für beide Fahrzeugtypen eine Vielzahl von objektiven Parametern. Diese umfassen Fahrparameter, wie z.B. die Geschwindigkeit oder die Querbewegung. Dazu gibt es Daten über die verschiedenen im Fahrzeug verbauten Systeme, wie z.B. *Adaptive Cruise Control (ACC)* oder *Lane Departure Warning (LDW)*. Daten von Sensoren, die diesen Systemen Informationen über Abstand zum Vorderfahrzeug oder zur Spurmarkierung liefern, sind ebenfalls vorhanden. Ebenfalls wurden GPS-Koordinaten während der Fahrt aufgezeichnet.

Die in euroFOT erhobenen Daten wurden nachträglich bei den Fahrzeugherstellern sowie am Würzburger Institut für Verkehrswissenschaften während des Projekts aufbereitet. Schwerpunkt lag dabei auf der Plausibilisierung und Interpolation der Daten über Fahrdynamik. Dort wurde auch eine Datenauswahl getroffen und die Anzahl der Variablen auf ca. 800 begrenzt.

Diese im Fahrzeug direkt aufgezeichneten Daten wurden mit Informationen aus der digitalen Karte sowie weiteren Parametern angereichert. Die Informationen aus der digitalen Karte wurden mit Hilfe eines externen Map-Matching-Tools extrahiert. Dazu wurden bereits erste Bereinigungsschritte der durch das externe Matchingverfahren ermittelten Informationen durchgeführt. Insgesamt steht somit pro Fahrt ein umfassender Satz an objektiven Verlaufsdaten zur Verfügung, der die Analyse von Fahrten oder auch verschiedenen Abschnitten ermöglicht.

4.1.2 Kennzahlen der Datenbasis

Insgesamt sind Daten von 115 Fahrern verfügbar, die geplant je drei Monate an dem Versuch teilnahmen. Die Fahrer waren vorwiegend Kunden der beteiligten Autohersteller. Tabelle 4.2 gibt einen Überblick über die insgesamt in der Datenbasis verfügbaren Daten (getrennt für die beiden verwendeten Fahrzeugtypen). Insgesamt betrachtet wurden in knapp 40000 Fahrten über eine Million Kilometer gefahren. Dabei wurden mehr als 15000 Fahrstunden aufgezeichnet.

	Fahrzeugtyp 1	Fahrzeugtyp 2	Gesamt
Anzahl Probanden	64	51	115
Anzahl Fahrten	24 274	15 429	39 703
Insgesamt gefahrene Kilometer	629 870	383 392	1 013 262
Aufgezeichnete Fahrstunden	9 108	6 021	15 129
Anzahl Versuchstage	4 909	3 799	8 708

Tabelle 4.2: Übersicht über die Datenbasis, Tabelle nach [64]

In Abbildung 4.2 sind die gefahrenen Routen für die gesamte Stichprobe geografisch visualisiert. Die meisten Fahrten fanden in Deutschland mit Schwerpunkt im Südwesten statt. Dazu gab es Fahrten ins europäische Ausland.

Zusätzlich zur Beschreibung der gesamten Stichprobe sind in Tabelle 4.3 die fahrerweisen Daten (Mittelwert sowie Minima und Maxima) dargestellt. Dabei zeigen sich deutliche Unterschiede zwischen den Fahrern. Im Schnitt fuhren Fahrer 133.5 Stunden und legten knapp 9000 km zurück. Allerdings gibt es einen Fahrer, der die dreifache Dauer unterwegs war und dabei auch das Dreifache der Strecke zurücklegte. Die Minima hingegen liegen mit 20.7 Stunden und 1660 km deutlich unter den durchschnittlichen Werten. Auch die Anzahl der Fahrten variiert zwischen 61 und 975 bei einem Mittelwert von 351 Fahrten.

Allerdings ist zu bedenken, dass einige Fahrer die Studie abbrachen und somit nicht Daten von drei Monaten erhoben wurden. Dazu konnte es während der Fahrt zu Datenverlusten kommen. Daher werden auch die mittlere Fahrtdauer und -strecke pro Fahrt dargestellt. Auch hier zeigen sich Unterschiede zwischen den Fahrern. Bei einem Durchschnittswert von 0.4 Stunden gibt es ein Fahrer mit einem doppelt und ein Fahrer mit einem halb so großen Wert. Bei der Fahrtstrecke ist die Streuung noch größer (Mittelwert: 26.8 km, Minimum: 8.1 km, Maximum: 84.2 km).

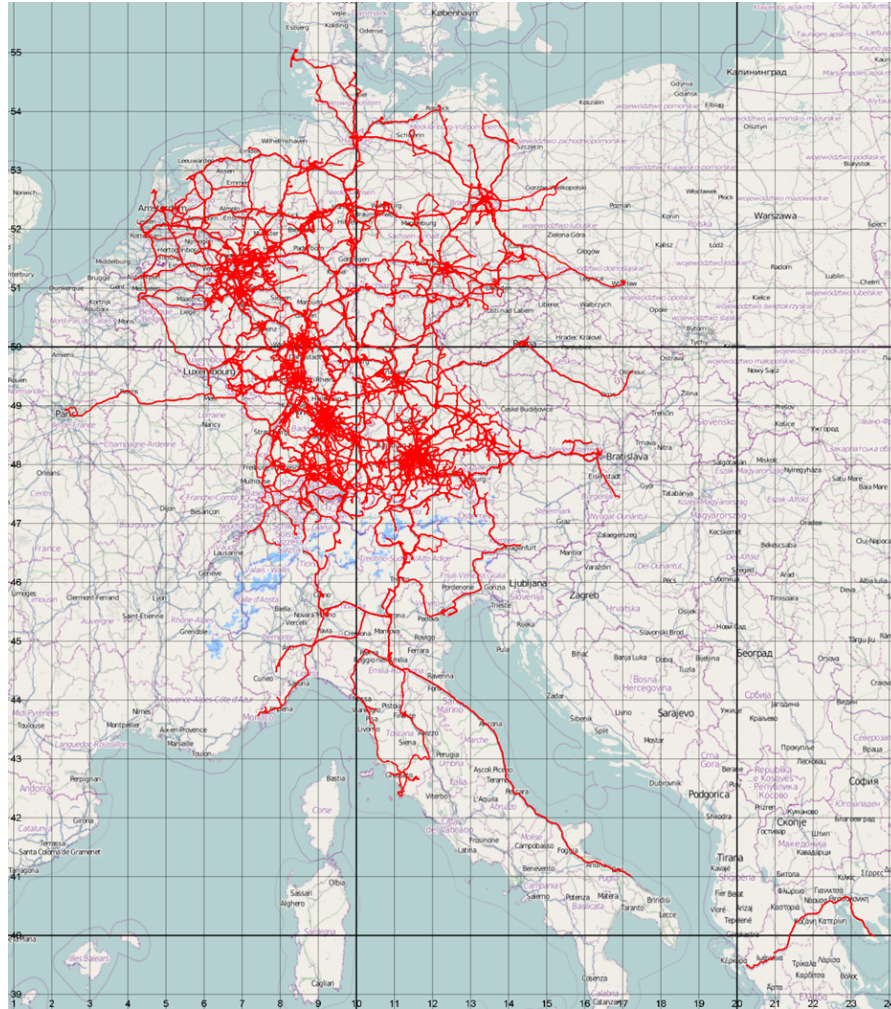


Abbildung 4.2: Grafische Darstellung aller aufgezeichneten Fahrten, Darstellung aus [64]

	mean	min	max
Fahrerweise Fahrtdauer [h]	133.5	20.7	394.5
Fahrerweise Fahrtstrecke [km]	8941	1660	27767
Anzahl Fahrten pro Fahrer [N]	351	61	975
mittlere Fahrtdauer [h]	0.40	0.16	0.88
mittlere Fahrtstrecke [km]	26.8	8.1	84.2

Tabelle 4.3: Beschreibung der erhobenen Daten im Untersuchungszeitraum pro Fahrer, Tabelle nach [64]

4.2 Einordnung der Datenbasis

Zusätzlich zu den objektiv erhobenen Daten für die im Untersuchungszeitraum durchgeführten Fahrten füllten alle Fahrer eine Reihe von Fragebögen zur Erhebung von *subjektiven Daten* aus. So waren nur vier der teilnehmenden Fahrer Frauen. Acht Fahrer gaben an, dass ihre Muttersprache nicht Deutsch ist. Im Durchschnitt waren die Fahrer 44 Jahre alt. (Minimum: 25 Jahre, Maximum: 66 Jahre). Angaben über die Fahrleistung im letzten Jahr reichten von 10 000 bis 140 000 Kilometer mit einem Mittelwert von ca. 44 400 Kilometern. Auch die Fahrpraxis variierte in Abhängigkeit des Alters stark zwischen 6 und 48 Jahren (Mittelwert: 25 Jahre). [64]

	mean	min	max
Alter [Jahre]	44	25	66
Fahrstrecke im letzten Jahr [km]	44 459	10 000	140 000
Fahrpraxis mit PKW [Jahre]	25	6	48

Tabelle 4.4: *Stichprobenbeschreibung von Basis der Fragebögen, Tabelle nach [64]*

Dazu wurden zu den allgemeinen Angaben vor allem folgende subjektive Daten erhoben [64]:

- Angaben zu Unfallbeteiligungen in den letzten Jahren.
- Angabe, ob und warum schon einmal der Führerschein entzogen wurde.
- *Sensation Seeking Questionnaire (SSQ)* nach Arnett [5].
- *Driver Behaviour Questionnaire (DBQ)* in einer Version mit 28 Items nach Mattsson [61].

„Der SSQ erfasst das Bedürfnis der Menschen, Neues bzw. Spannendes zu erleben. Für den Sensation Seeking Questionnaire lässt sich ein Gesamtwert, sowie Werte auf den Subskalen Novelty und Intensity auswerten. Es wird angenommen, dass Fahrer, die hohe Werte im Sensation Seeking aufweisen, auch im Fahren intensives Erleben (bspw. durch schnelles Fahren) suchen und deswegen zu einer riskanteren Fahrweise neigen.“[64]

„Der DBQ möchte die Tendenz der Fahrer zu Fahrfehlern bzw. Fahrregelverstößen erfassen. Er besteht aus den Subskalen Aggressive Violation, Ordinary Violation, Slips und Lapses. Die ersten beiden Skalen erfassen die Neigung zu Fahrregel-

verstößen, die beiden anderen zu Fahrfehlern.“[64] Die Auswertungen in Metz et al. [64] zu Korrelationen zeigten keinen Zusammenhang zwischen SSQ und DBQ.

4.2.1 Fahrertypen

Die im Projekt euroFOT auf unterschiedliche Weise erhobenen Daten werden in der Literatur häufig verwendet, um die Vorhersagekraft von Persönlichkeits- bzw. Fahrverhaltensfragebögen zu validieren [96]. Ebenso wird versucht, Fahrertypen zu definieren. Ein Ansatz ist es, bestimmte Eigenschaften der Fahrer (z.B. Persönlichkeit, subjektive Angaben zum Fahrverhalten) mit objektiven Kriterien auf den Ebenen der Bahnführung und Stabilisierung in Verbindung zu setzen. [64]

In den Untersuchungen von Metz et al. [64] hierzu zeigte sich, dass zwischen den erhobenen Parametern zu Unfällen und der Subskala Ordinary Violation des DBQ signifikante Korrelationen zu finden sind. Fahrpraxis und Fahrleistung standen dabei in keinem Zusammenhang zu den berichteten Unfällen. Die Angaben des SSQ hingegen zeigten keine signifikanten Zusammenhänge zum Fahrverhalten.

Neben den Zusammenhängen zwischen den verschiedenen subjektiven Indikatoren des Fahrverhaltens, wurden in Metz et al. [64] auch Zusammenhänge zwischen objektiven Fahrverhaltensparametern untersucht. So konnte z.B. gezeigt werden, dass eine höhere mittlere Geschwindigkeit im Zusammenhang mit einem höheren Anteil an Geschwindigkeitsübertretungen steht. Ebenso steht innerorts und auf der Autobahn schnelleres Fahren mit geringeren präferierten Abständen in Zusammenhang.

„Somit ergeben sich zusammenfassend in den objektiven Parametern Hinweise auf Präferenzen im Fahrstil auf den Ebenen der Bahnführung und Stabilisierung. Es gibt Fahrer, die einen eher offensiven Fahrstil und Fahrer, die einen eher defensiven Fahrstil bevorzugen. Dies zeigt sich sowohl in der gewählten Geschwindigkeit als auch in den präferierten Abständen. Fahrer, die einen offensiveren Fahrstil zeigen, neigen auch stärker zu Fahrregelverstößen (gemessen über sehr kleine Abstände und Geschwindigkeitsüberschreitungen). Dies lässt sich auch im DBQ abbilden und zwar erwartungsgemäß in den Skalen zu Fahrregelverstößen. Innerorts steht außerdem eine offensivere Fahrweise (gemessen über objektive Fahrregelverstöße) mit der Unfallhäufigkeit in Verbindung.“[64]

Die Ergebnisse zeigen, dass es in der Stichprobe unterschiedliche Fahrer gibt. Inwieweit es einen Zusammenhang zwischen dem Verhalten der Fahrer auf der Ebene der Navigation und den Resultaten für die unteren Ebenen der Bahnführung und Stabilisierung gibt, wäre zu überprüfen.

4.2.2 Mobilität in Deutschland 2008

„Für eine Einordnung der Fahrer in Hinblick auf ihre Mobilität wurden die entsprechenden Kennzahlen mit Werten der Studie *Mobilität in Deutschland 2008 (MiD2008)* [33] verglichen (siehe Tabelle 4.5). In dieser Studie wurde eine große repräsentative Stichprobe über ihr Mobilitätsverhalten an einem Stichtag befragt. Daraus wurden Kennzahlen für die Mobilität in ganz Deutschland abgeleitet. Dies erfolgte anhand der Anzahl Wege pro Tag sowie der Wegdauer und -länge pro Fahrt. Zur besseren Einschätzung der Größen werden zusätzlich Minimum und Maximum der gemittelten Werte pro Fahrer angegeben.“[64]

	MiD2008	euroFOT-Stichprobe		
		mean	min	max
Anzahl Wege als Fahrer pro Tag	1.46	4.75	1.61	7.43
Dauer pro Weg [min]	21.0	24.0	9.6	52.8
Strecke pro Weg [km]	14.7	26.8	8.1	84.2

Tabelle 4.5: Vergleich der euroFOT-Stichprobe mit der MiD2008-Studie, Tabelle nach [64]

„Die Fahrer der euroFOT-Stichprobe führen deutlich mehr Fahrten pro Tag durch als es in der MiD2008 als repräsentativ angegeben wird. Ebenso unterscheidet sich die Wegstrecke in gleicher Richtung, wohingegen die Wegdauer pro Fahrt vergleichbar bleibt. Dies lässt auf einen höheren Anteil an Autobahnfahrten schließen. Die untersuchte Stichprobe besteht somit überwiegend aus Häufig- und Langfahrern. Was die Repräsentativität der Ergebnisse weiter einschränkt ist, dass die Stichprobe fast ausschließlich aus Männern mittleren Alters besteht.“[64]

Damit ist offensichtlich, dass die Stichprobe nicht für die Gesamtheit der Fahrer in Deutschland repräsentativ ist. Ergebnisse können nicht, wie z.B. für die MiD2008, auf ganz Deutschland übertragen werden. Allerdings zeigen sich auch innerhalb der Stichprobe für alle objektiven sowie subjektiven Parameter Unterschiede zwischen den Fahrern. Ein weiterer Vorteil der Stichprobe ist dazu die Vertrautheit mit den von den Herstellern bereitgestellten Fahrzeugen sowie die starke Nutzung der Fahrzeuge über die drei Monate. Damit entsteht eine sehr große Datenbasis ohne Anpassungseffekte beim Verhalten, die für die Untersuchungen und die Entwicklung eines Verfahrens zur Analyse des Routenauswahlverhaltens genutzt werden kann.

5 Verfahren zur modellgestützten Personalisierung

5.1 Schritte des Verfahrens

In diesem Abschnitt wird das an einen KDP angelehnte Verfahren zur modellgestützten Personalisierung von Navigationssystemen detailliert beschrieben und erläutert. Der Schwerpunkt liegt auf den einzelnen Schritten der Modellierung (Abschnitte 5.2 und 5.3), der Personalisierung (Abschnitt 5.4) und der Bewertung (Abschnitt 5.5). Wie bereits in Kapitel 3 dargestellt, werden verschiedene Perspektiven und Dimensionen in die Analyse der über den Fahrer erhobenen Daten einbezogen. In Abbildung 5.1 sind die einzelnen Komponenten des Verfahrens, ihre Abhängigkeiten sowie ihre Zugehörigkeit zu den einzelnen Schritten aus Abbildung 3.10 dargestellt.

Ausgangspunkt sind die in euroFOT erhobenen Realfahrdaten, die zunächst aufbereitet werden. Die Datenaufbereitung (weiß) verfolgt das Ziel die Datensätze für die drei Perspektiven sowie für die Kartenerstellung zu generieren und umfasst die Datenauswahl, -bereinigung und -reduktion. Die Modellierung (hellgrau) umfasst die Analysen der Kanten-, Routen- und Netzwerkperspektive. Für die Modelle der Perspektiven wird eine Korrelationsanalyse durchgeführt und darauf aufbauend ein zweidimensionales Modell erstellt.

Auf den Ergebnissen der Modellierung baut die Personalisierung (grau) auf. Die drei Ansätze zur Personalisierung der Kantengewichte, der Zielfunktion und einer kombinierten Personalisierung werden mit Hilfe einer internen Routingreferenz implementiert und bewertet (dunkelgrau). Das Ergebnis des gesamten Verfahrens ist eine bewertete personalisierte Route.

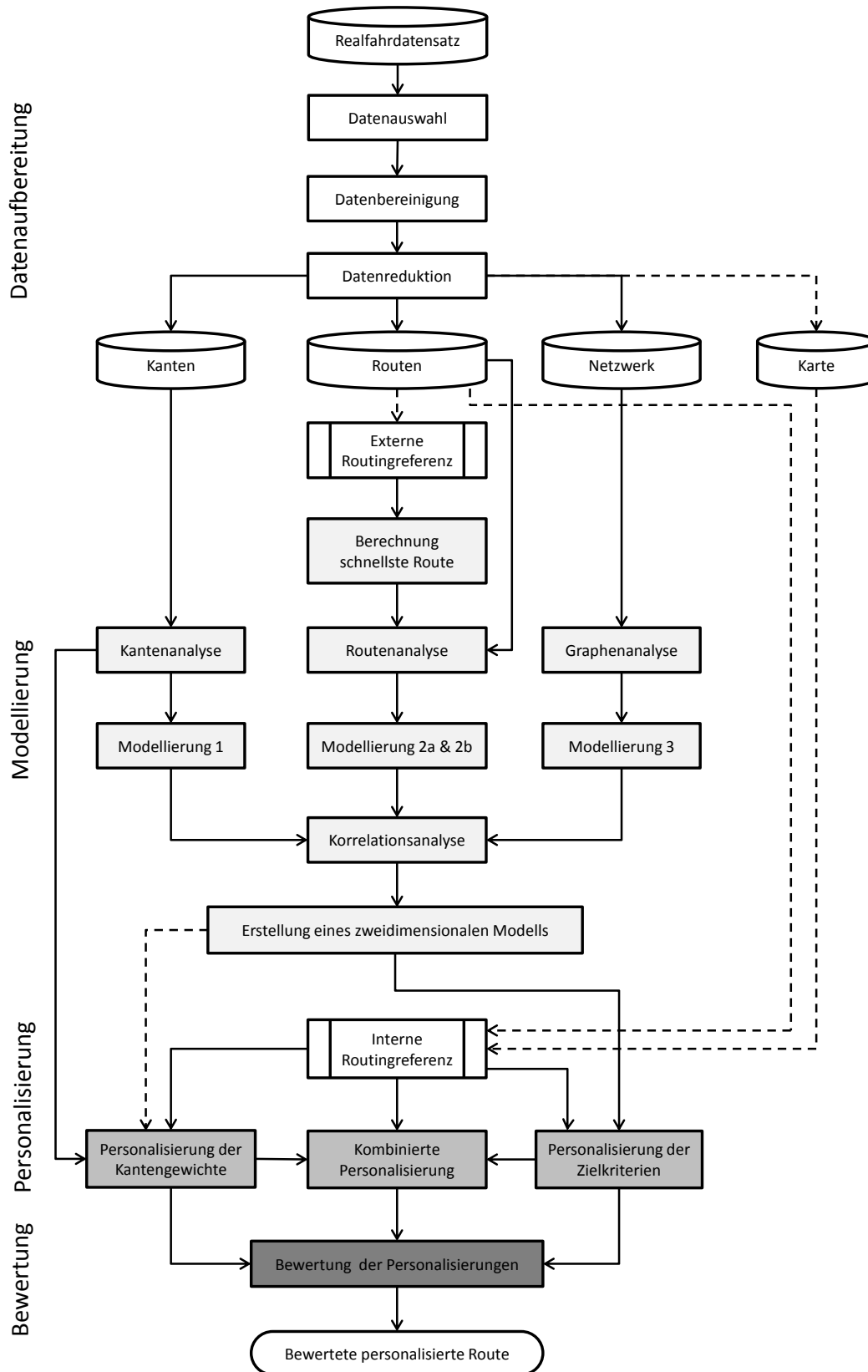


Abbildung 5.1: Aufbau des Verfahrens

5.1.1 Datenaufbereitung

In diesem Abschnitt werden das Vorgehen und die Anforderungen an die Datenaufbereitung detaillierter beschrieben. Die Umsetzung für die Realfahrdaten des euroFOT-Projekts sind in Kapitel 6 dargestellt.

Datenauswahl

Zunächst werden die Fahrer ausgewählt, für die über einen längeren Zeitraum Fahrten erhoben wurden. Als nächstes muss geprüft werden, welche Fahrten und welche Abschnitte davon für eine Analyse des Routenauswahlverhaltens der Fahrer genutzt werden können. So kann es vorkommen, dass aufgrund von Ausfällen der Messtechnik Fahrten nicht korrekt aufgezeichnet wurden. Diese Fahrten sind auszuschließen.

Abschließend werden die relevanten Variablen für jede Fahrt ausgewählt. Auch für diese muss geprüft werden, dass es zu keinem Ausfall der Messtechnik für eine relevante Variable gekommen ist. Das Ergebnis der Datenauswahl ist eine Menge von Fahrten von Fahrern, die die Grundlage für das weitere Vorgehen bilden.

Datenbereinigung

Da die Daten im Realverkehr erhoben werden, kann es bei den ausgewählten Variablen zu kurzzeitigen Datenaussetzern kommen. Zwar werden die Kriterien der Datenauswahl erfüllt, doch ist eine Datenbereinigung notwendig, damit die Qualität der Daten der folgenden Analysen genügt. Signale über die Fahrzeugdynamik, wie z.B. die Geschwindigkeit, müssen geprüft und unter Umständen interpoliert werden. Dazu ist es notwendig, weitere Daten aus den ursprünglich erhobenen abzuleiten.

Ein Hauptaugenmerk wird auf die erhobenen Positionsdaten gelegt. Sie sind Kern der meisten Analyseschritte und bedürfen daher einer gründlichen Kontrolle, zumal das zugrunde liegende GPS-Signal vielen Störeinflüssen ausgesetzt ist (siehe Abschnitt 2.2.3).

Dazu muss das Map Matching korrekte Informationen der digitalen Karte zu den aufgezeichneten Positionen liefern. Daher ist es notwendig, diese Informationen ebenfalls zu bereinigen. Das Ergebnis sind gesicherte und bereinigte Positionsdaten und die dazugehörigen Informationen.

Datenreduktion

Ziel der Datenreduktion ist die Bereitstellung von drei verschiedenen Datensätzen für die jeweiligen Analyseperspektiven sowie einen für die Erstellung einer digitalen Karte. Für die Kantenperspektive müssen aus den erhobenen Daten Kanten und

deren Charakteristika identifiziert werden. Eine Route kann als Kantenzug angesehen werden. Jede Kante als Teil der Route besitzt Charakteristika der digitalen Karte, z.B. eine ID oder die erlaubte Geschwindigkeit. Dazu beinhaltet sie auch gemessene Charakteristika auf der Kante, wie z.B. die gefahrene Geschwindigkeit des Fahrers. Für die folgenden Analysen werden die Zeit, die ein Fahrer auf einer Kante verbracht hat (*Messdauer*), sowie die Länge der auf der Kante gefahrenen Strecke (*Messlänge*) benötigt. Wurde eine Kante mehrfach benutzt, so kommen diese Parameter für die Kante entsprechend oft vor.

Für die Untersuchungen der Routenperspektive werden die gesamten Routen mit Hilfe einer externen Routingreferenz analysiert. Für die Routingreferenz werden Start- und Zielposition und für die Analyse sowohl die Routendauer wie auch die Routenlänge benötigt. Der Routenverlauf und die entsprechenden Positionen der gemessenen Route werden ebenfalls aufgenommen, um sie für die Personalisierung der Routendauer zu nutzen.

Die Routen werden unter Berücksichtigung zweier Charakteristika unterteilt. Die Charakteristika betreffen die Bekanntheit einer Route sowie die Nutzung des Navigationssystems. Führt der Fahrer eine bekannte Route ohne Nutzung des Navigationssystems, so ist anzunehmen, dass der Fahrer die Route gemäß seiner Präferenzen ausgewählt und auch eine entsprechende Fahrweise hat. Daher sind diese Fahrer-Realrouten bzw. *FR-Routen* für die Analysen des Fahrerverhaltens nutzbar.

	Bekannte Route	Unbekannte Route
Navigation nicht genutzt	Fahrer-Realrouten	Fahrer-Experimentalrouten
Navigation genutzt	Navigations-Zusatzrouten	Navigations-Anwendungsrouten

Tabelle 5.1: *Unterteilung der Routen*

Wird das Navigationssystem genutzt, so lässt sich der Fahrer führen und übergibt damit seine Routenauswahlkompetenz an das Navigationssystem. Im unbekannten Gebiet erfüllt das Navigationssystem seine Aufgabe, den Fahrer entlang einer ihm unbekannten Route zu führen (*Navigations-Anwendungsrouten*). Wird das Navigationssystem auf einer dem Fahrer bekannten Route genutzt, sind die Gründe nicht eindeutig. Das System kann zur Unterstützung, als Alternativvorschlag aber auch als Informationssystem über die aktuelle Verkehrslage genutzt werden. Dementsprechend erfüllt das Navigationssystem in diesem Fall eine Zusatzfunktion (*Navigations-Zusatzrouten*). Nutzt der Fahrer das Navigationssystem nicht, obwohl ihm die Route unbekannt ist (*Fahrer-Experimentalrouten*), werden diese Routen in der weiteren Analyse nicht berücksichtigt. Eine Einbeziehung

könnte die Ergebnisse verzerren, da die Gründe für die Nichtnutzung unbekannt sind.

Die Analysen der Netzwerkperspektive umfassen die Gesamtheit aller gemessenen Kanten eines Fahrers. Grundlage bildet der Kantenzug einer Route. Wie für die Kantenperspektive werden die Charakteristika aus der digitalen Karte, insbesondere die ID, hinzugefügt. Dazu werden die Positionen vom Anfang und Ende der Kante gespeichert. Das Ergebnis ist eine Kantenmenge inklusive der Positionen für jeden Fahrer. Daraus wird ein knotenbasierter Graph generiert. In Kapitel 6 werden die Datenreduktionsschritte für den genutzten Realfahrdatensatz veranschaulicht.

Der Erstellung des Datensatzes für die digitale Karte erfolgt auf Basis der genutzten Kanten und ihrer logischen Zusammenhänge. Hierfür werden die Daten aller Fahrer genutzt, um eine möglichst große Karte zu erzeugen, die durch einen kantenbasierten Graphen repräsentiert wird.

5.1.2 Analyse und Modellierung

Zunächst werden die Daten separat für die drei Perspektiven analysiert. Unabhängig voneinander werden Kennwerte für verschiedene Maße, die die navigatorischen Präferenzen eines Fahrers charakterisieren, für die Dimensionen der Zeit und für die Dimension des Raumes bestimmt. Der Ablauf der einzelnen Analysen wird in Abschnitt 5.2 detailliert beschrieben.

Wie in Tabelle 3.1 dargestellt, ist das Ergebnis der Kantenanalyse ein Modell der Dimension der Zeit (Modell 1). Hierbei werden die gemessenen Kantengewichte der Geschwindigkeit mit denen in der digitalen Karte gespeicherten verglichen. Die Analyse der Routen untersucht sowohl die Dimension der Zeit wie auch die Dimension des Raumes und stellt damit einen Bezug zwischen beiden her. Dabei werden die gemessenen Routen mit den berechneten Routen der externen Routingreferenz verglichen. Das Ergebnis sind zwei Modelle für die beiden Dimensionen zur Charakterisierung der Routen (Modell 2a und 2b).

Für die Netzwerkperspektive, die das gesamte Mobilitätsverhalten eines Fahrers umfasst, wird die graphentheoretische Repräsentation der Gesamtheit der gemessenen Routen untersucht. Dies ergibt einen aggregierten Kennwert, der zur Dimension des Raumes gehört (Modell 3). Damit stellen die drei Perspektiven den Schritt der Datentransformation dar, bei dem Daten in Informationen umgewandelt werden.

Um die generierten Informationen für die Routenberechnung nutzbar zu machen, müssen diese im zweiten Schritt der Modellierung miteinander in Zusammenhang gesetzt und ein zweidimensionales Modell erstellt werden. Dazu wird zunächst eine *Korrelationsanalyse* durchgeführt und Beziehungen der Modelle in Abhängigkeit der Kennwerte ermittelt. Die Beschreibung ist in Abschnitt 5.3 zu finden. Dies bildet die Grundlage für ein zweidimensionales Modell der navigatorischen Präferenzen, das als Wissen für eine Personalisierung genutzt werden kann.

5.1.3 Personalisierung und Bewertung

Die Personalisierung besteht aus verschiedenen aufeinander aufbauenden Ansätzen. Dabei werden zunächst eine Personalisierung der Kantengewichte [53] sowie eine Erstellung einer personalisierten Zielfunktion [53] durchgeführt. Dazu werden diese Ansätze miteinander kombiniert. Die Personalisierung der Kantengewichte stellt eine Anpassung der Grundlage der Routenberechnung dar. Wie bereits in Abschnitt 2.3.5 beschrieben, ist dieses Verfahren in seiner rudimentären Form bereits in heutigen Navigationssystemen implementiert. Analog dazu wird eine Personalisierung durchgeführt.

Ebenso wird eine zweidimensionale Zielfunktion erstellt. Die zugrunde liegende Modellierung umfasst sowohl die Dimension der Zeit wie auch die Dimension des Raumes. Dementsprechend umfasst die Zielfunktion diese beiden Dimensionen. Die kombinierte Personalisierung verbindet die beiden Ansätze, um so das gesamte Potenzial der Personalisierung auszuschöpfen. Eine detaillierte Beschreibung ist in Abschnitt 5.4 zu finden.

Abschließend wird die Personalisierung bewertet. Ziel der Personalisierung ist es, eine Route zu berechnen, die die Präferenzen eines Fahrers in größerem Maße berücksichtigt als eine Standardroute. Für die Bewertung werden drei verschiedene Gütemaße genutzt. Das erste Maß der Zielkriterien betrifft die Routendauer sowie die Routenlänge der personalisierten Routen. Ein weiteres Gütemaß analysiert die Anteile der Straßenklassen im Vergleich zu den Anteilen der FR-Routen (Gütemaß der Streckencharakteristika). Für das dritte Gütemaß wird die Übereinstimmung zwischen der FR-Route und den berechneten Routen analysiert. Die verschiedenen Gütemaße werden in Abschnitt 5.5 detailliert beschrieben.

5.2 Analyse der navigatorischen Präferenzen der Fahrer

5.2.1 Kantenperspektive

Der Ansatz für die Analysen der Kantenperspektive orientiert sich an existierenden Verfahren heutiger Navigationssysteme, siehe Abschnitt 2.3.5. Nach der Aufbereitung und Reduktion der Kanten liegen die von jedem Fahrer auf den verschiedenen Routen genutzten Kanten k als Kantenmenge vor. Diese ist unabhängig von der ID, so dass mehrere Kanten mit der gleichen ID in der Kantenmenge vorkommen können, hier aber als mehrere Kanten behandelt werden. Jeder Kante sind Charakteristika der digitalen Karte (z.B. Länge, angenommene Geschwindigkeit) sowie die gemessenen Charakteristika (Dauer, Länge) zugeordnet. Auf Basis der gemessenen Charakteristika wird für jede Kante k eines Fahrers f eine gemessene Geschwindigkeit $v_{k,f}(obs)$ berechnet.

Der Aufbau der Analyse für die Kanten eines Fahrers f ist in Abbildung 5.2 dargestellt. Die Verbindung zu anderen Komponenten des Verfahrens sind gestrichelt dargestellt. Ausgangspunkt der Kantenanalyse sind die in der Datenreduktion erstellten Kantenmengen für jeden Fahrer. Für jede Kante k eines Fahrers f wird eine gemessene Geschwindigkeit $v_{k,f}(obs)$ auf Basis der Messlänge und der Messdauer berechnet. Dazu werden die Messungen einer Geschwindigkeitsklasse g entsprechend der vorgegebenen Geschwindigkeit $v_g(map)$ für das Routing aus der digitalen Karte zugeordnet.

Für jeden Fahrer f wird pro Geschwindigkeitsklasse g ein neues Kantengewicht der Geschwindigkeit $v_{g,f}(obs)$ berechnet. Dafür wird der Mittelwert für die Menge der gemessenen Geschwindigkeiten der Kanten der Kantenklasse $V_{g,f}$ gebildet, siehe Gleichung 5.1. Das Ergebnis sind *Individualkantengewichte* für jeden Fahrer und alle Geschwindigkeitsklassen, die für eine personalisierte Routenberechnung genutzt werden können.

$$v_{g,f}(obs) = mean(V_{g,f}) \quad \forall \text{ Klassen } g \text{ und Fahrer } f \quad (5.1)$$

Für die Erstellung des Modells 1 werden im ersten Schritt die Individualkantengewichte für jede Klasse mit den in der digitalen Karte hinterlegten Gewichten verglichen. Dabei wird ein *0-Relativ-Vergleich* durchgeführt. Hierfür wird eine Fallunterscheidung durchgeführt, um den Effekt von größeren Abweichungen in beide

Richtungen gleich zu gewichten. Entsprechend wird ein Wert $rv_{g,f}$ für die Abweichung einer Klasse g bestimmt, siehe Gleichung 5.2.

$$rv_{g,f} = \begin{cases} \frac{v_{g,f}(obs)}{v_g(map)} - 1, & \text{if } v_{g,f}(obs) \leq v_g(map) \\ 1 - \frac{1}{\frac{v_{g,f}(obs)}{v_g(map)}}, & \text{if } v_{g,f}(obs) > v_g(map) \end{cases} \quad \forall \text{ Klassen } g \text{ und Fahrer } f \quad (5.2)$$

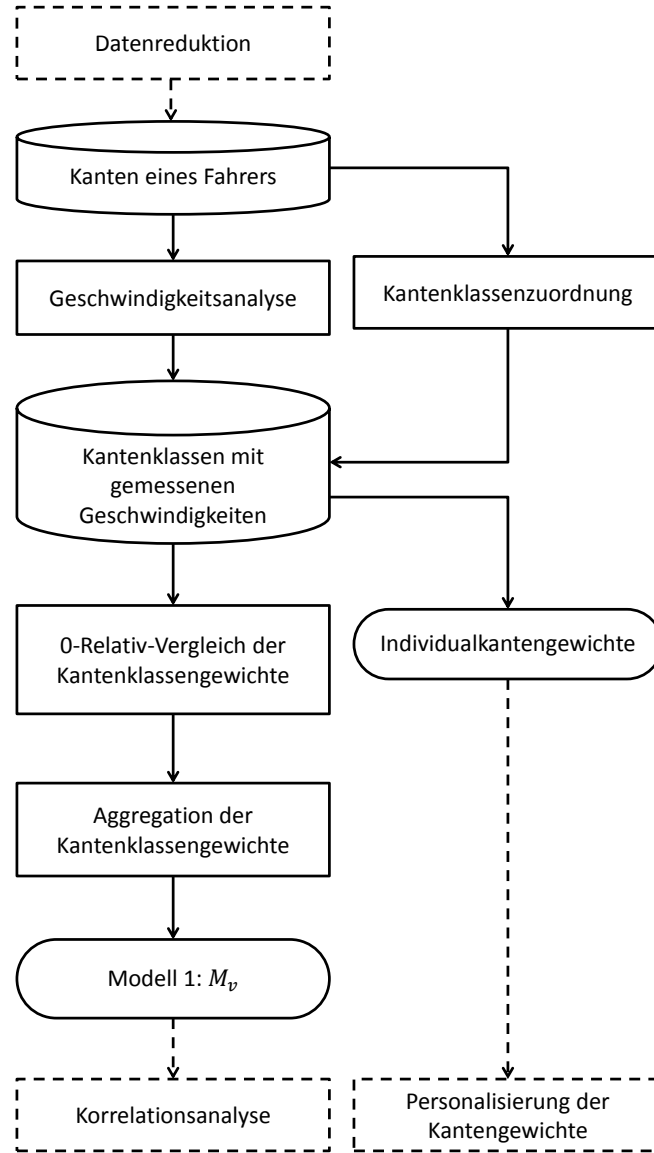


Abbildung 5.2: Ablauf der Analyse für die Kantenperspektive

Stimmen beide Werte $v_{g,f}(obs)$ und $v_g(map)$ überein, so ergibt sich ein Wert von Null. Ist $v_{g,f}(obs)$ kleiner als $v_g(map)$, so ergibt sich ein negativer Wert zwischen -1

und 0, der die Größe der Abweichung angibt. In diesem Fall ist der Fahrer langsamer als in der digitalen Karte gespeichert. Für den Fall, dass $v_{g,f}(obs)$ größer ist als $v_g(map)$ und der Fahrer damit schneller ist als in der digitalen Karte gespeichert, ergibt sich ein positiver Wert zwischen 0 und 1.

Dieses Vorgehen hat zur Folge, dass große Abweichungen abgeschwächt werden, um einen zu großen Einfluss dieser Werte zu vermeiden. Dies kann bei gemessenen Geschwindigkeiten vorkommen und lässt sich am Beispiel eines beobachteten Wertes, der einmal doppelt und einmal nur halb so groß ist, verdeutlichen.

Bei einem einfachen relativen Vergleich würde einmal ein Wert von 2 und einmal ein Wert von 0.5 berechnet. Dies würde einen Mittelwert von 1.25 ergeben. Im Vergleich zur Gleichheit mit dem Wert 1 würde dies ein Ungleichgewicht nach sich ziehen und zudem dafür sorgen, dass geringere Abweichungen unterrepräsentiert werden.

Für das in dieser Arbeit gewählte Verfahren ergibt sich für das Beispiel Werte von -0.5 und 0.5. Damit liegt der Mittelwert bei Null und die Auswirkung der beiden Werte heben sich gegenseitig auf. Ebenso werden Auswirkungen auf die folgende Aggregation abgeschwächt. Insbesondere für noch größere Abweichungen ist dies wichtig, da der Wert für das gewählte Verfahren nicht unter bzw. über -1 bzw. 1 fallen kann. Damit unterliegt dieses Maß einer gewissen Trägheit.

Für die Modellierung und das entsprechende Maß M_v der Kantenperspektive wird der Mittelwert über die Menge der Werte der Geschwindigkeitsklassen RV_f gebildet und zu einem Kennwert aggregiert. Dabei fließen alle validen Werte ein, die für Werte auf Basis einer genügend großen Anzahl an Kanten ermittelt haben, um für jeden Fahrer f einen Wert M_{v_f} zu erhalten, siehe Gleichung 5.3.

$$M_{v_f} = \text{mean}(RV_f) \quad \forall \text{ Fahrer } f \quad (5.3)$$

Das Ergebnis ist ein Maß M_{v_f} für jeden Fahrer bezüglich der Dimension Zeit, siehe Gleichung 5.3, das zusammen mit den anderen Maßen in die Korrelationsanalyse und die zweidimensionale Modellierung eingeht. Dazu werden die Individualkantengewichte für die Personalisierung der Kantengewichte genutzt.

5.2.2 Routenperspektive

Für die Modellierung wird für die Start- und Zielpositionen der FR-Routen mit einer Mindestlänge von 1 km eine schnellste Route mit Hilfe einer externen Routingreferenz berechnet. Die externe Routingreferenz berechnet Routen für eine beliebige Start-Zielkombination und einer gegebenen Zielfunktion. Für die Eingabe wird

in dieser Arbeit ein *Application Programming Interface* (API) [10] genutzt, um die Routenberechnung mit Hilfe der Routingfunktion von Google durchzuführen. Start- und Zielpunkte sowie potentielle Zwischenziele sind als GPS-Koordinaten in einer separaten Datei gespeichert und werden von der API aufgerufen. In der API selbst kann die Zielfunktion, Restriktionen (z.B. Autobahn vermeiden) sowie die Art der Fortbewegung (z.B. Fahrzeug oder Fußgänger) angegeben werden.

Auf Basis der Eingaben berechnet die externe Routenberechnung eine Route entsprechend des Zielkriteriums, falls diese existiert. Die Basis bildet die bei Google gespeicherte digitale Karte. Sowohl Karte wie auch das Routingverfahren sind eine Black Box für den Anwender der Referenz. Das Ergebnis ist eine Ausgabedatei, die sowohl Routendauer wie auch Routenlänge beinhaltet. Dazu werden die GPS-Koordinaten des Routenverlaufs in der Datei mit ausgegeben. Eine vereinfachte Darstellung des Aufbaus der externen Routingreferenz ist in Abbildung 5.3 gegeben.

Mit Hilfe der externen Routingreferenz ist es möglich, eine Vielzahl von Routen in kurzer Zeit zu berechnen und Routencharakteristika als Ergebnis zu erhalten. Die Eingabe von Zwischenzielen ermöglicht eine eingeschränkte Manipulation der Routenberechnung. Allerdings ist es nicht möglich, direkten Einfluss auf die digitale Karte als Grundlage der Routenberechnung oder der Berechnung selbst zu nehmen.

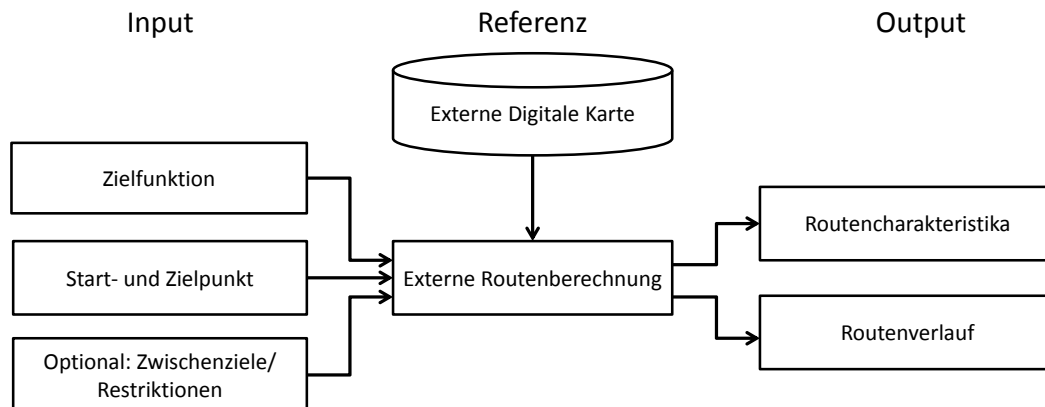


Abbildung 5.3: *Struktur der externen Routingreferenz*

In Abbildung 5.4 ist der Ablauf der Routenperspektivenanalyse dargestellt. Für die Berechnung der schnellsten Route werden keine Zwischenziele genutzt sondern ausschließlich die extrahierten Start- und Zielpositionen der FR-Routen r . Die Charakteristika der Routendauer $d_r(calc)$ und der Routenlänge $l_r(calc)$ werden extra-

hiert und mit den Routencharakteristika $d_{r,f}(obs)$ und $l_{r,f}(obs)$ der entsprechenden Fahrer-Realrouten verglichen.

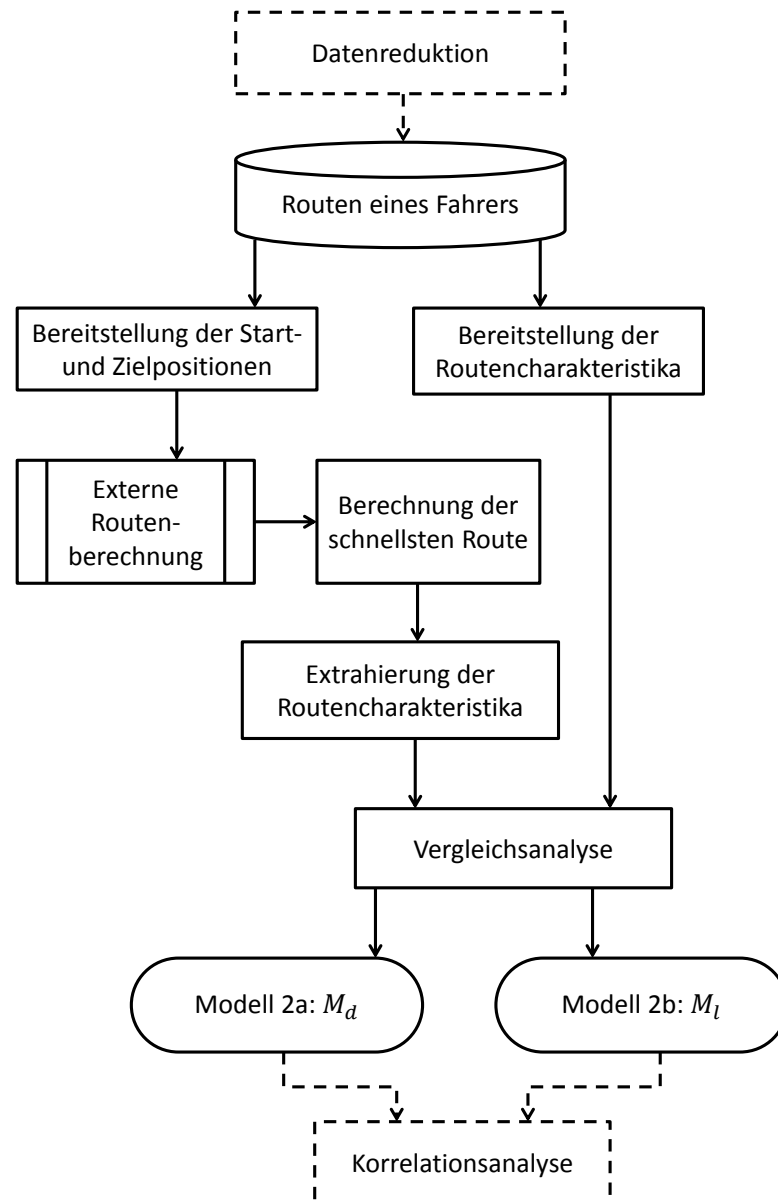


Abbildung 5.4: *Ablauf der Analyse für die Routenperspektive*

Das Vorgehen zur Bestimmung der beiden Maße M_d und M_l für die Modelle 2a und 2b verläuft analog zu der Berechnung der Werte für Geschwindigkeitsklassen. Der 0-Relativ-Vergleich zeigt, ob der Fahrer eine kürzere oder längere Route gewählt bzw. ob er dafür mehr oder weniger Zeit benötigt hat als mit Hilfe der externen

Routingreferenz berechnet. Die Formel für die entsprechenden Werte $rd_{r,f}$ und $rl_{r,f}$ sind in den Gleichungen 5.4 und 5.5 zu finden.

$$rd_{r,f} = \begin{cases} \frac{d_{r,f}(obs)}{d_r(calc)} - 1, & \text{if } d_{r,f}(obs) \leq d_r(calc) \\ 1 - \frac{1}{\frac{d_{r,f}(obs)}{d_r(calc)}}, & \text{if } d_{r,f}(obs) > d_r(calc) \end{cases} \quad \forall \text{ Routen } r \text{ und Fahrer } f \quad (5.4)$$

$$rl_{r,f} = \begin{cases} \frac{l_{r,f}(obs)}{l_r(calc)} - 1, & \text{if } l_{r,f}(obs) \leq l_r(calc) \\ 1 - \frac{1}{\frac{l_{r,f}(obs)}{l_r(calc)}}, & \text{if } l_{r,f}(obs) > l_r(calc) \end{cases} \quad \forall \text{ Routen } r \text{ und Fahrer } f \quad (5.5)$$

Ist die gemessene Routendauer kürzer oder länger als die Routendauer der berechneten schnellsten Route und damit der Wert negativ bzw. positiv, so stimmen die Annahmen für die Kantengewichte der Geschwindigkeit als Basis der Berechnung (schnellste Route) nicht mit der tatsächlich für den Fahrer gemessenen Geschwindigkeiten überein. Sind längere gemessene Routendauern durch externe Einflussfaktoren, wie z.B. Stau, erklärbar, so müssen kürzere Dauern durch das Verhalten des Fahrers verursacht sein.

Für die Bildung der Kennwerte M_{d_f} und M_{l_f} für jeden Fahrer f wird nun der Median über die Menge der Vergleichswerte RD_f und RL_f für alle Fahrer f gebildet und das Vorzeichen geändert, siehe Gleichungen 5.6 und 5.7.

$$M_{d_f} = -\text{median}(RD_f) \quad \forall \text{ Fahrer } f \quad (5.6)$$

$$M_{l_f} = -\text{median}(RL_f) \quad \forall \text{ Fahrer } f \quad (5.7)$$

Da es deutlich weniger Routen als Kanten gibt, wird an dieser Stelle der Median als robustes Maß genutzt, um den Einfluss etwaiger Ausreißer zu verringern. Das Vorzeichen wird geändert, um analoge Werte zu den Ergebnissen der Kantenperspektive zu bekommen. Ist der Fahrer besser, d.h. schneller oder wählt eine kürzere Route als die Referenzroute, so ist der entsprechende Wert positiv. Ist der Fahrer schlechter, so hat das einen negativen Wert zur Folge. In der Korrelationsanalyse in Abschnitt 5.3.1, in die die berechneten Maße eingehen, wird darauf noch einmal eingegangen.

5.2.3 Netzwerkperspektive

In diesem Abschnitt wird die Gesamtheit der gemessenen Routen mit Hilfe von graphentheoretischen Analysewerkzeugen untersucht und der Fahrer einem *Navigationstyp* [51] zugeordnet, der sein Mobilitätsverhalten beschreibt. Daraus wird

das Maß der Abdeckung M_a für das Modell 3 bestimmt. In Abbildung 5.5 ist der Ablauf der Analyse dargestellt.

Grundlage ist der in der Datenaufbereitung für jeden Fahrer erstellte Graph, der das Netzwerk des Fahrers repräsentiert. Dieser besteht aus Knoten und ungerichteten Kanten. Das Kantengewicht wird durch die Häufigkeit der Benutzung bestimmt. Der Schwerpunkt liegt auf dem Zusammenhang des Graphen.

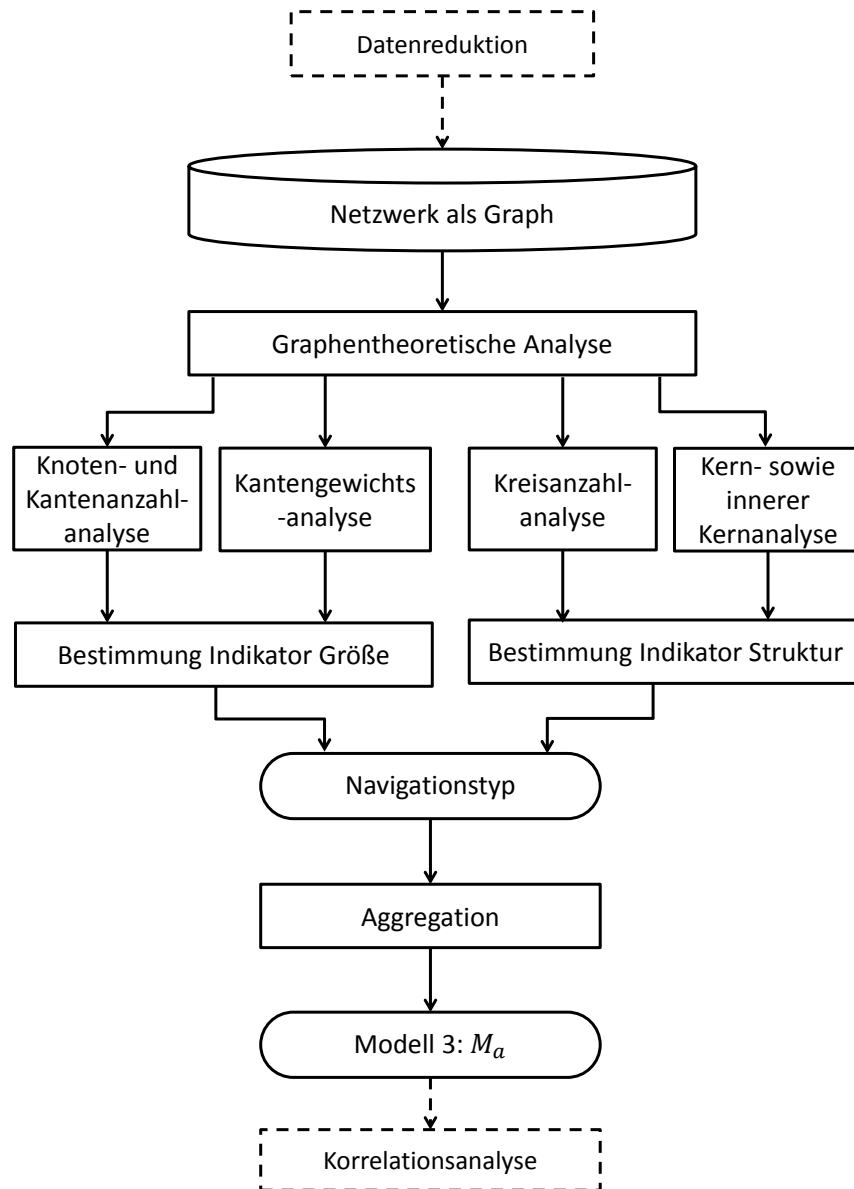


Abbildung 5.5: Ablauf der Analyse für die Netzwerkperspektive

Es werden zwei Indikatoren (*Größe* und *Struktur*) für den Graphen bestimmt, die das Mobilitätsverhalten des Fahrers quantifizieren sollen. Der Wert für den Indikator $I_{Größe}$ des Graphen wird durch die Anzahl der Knoten p und Kanten q

bestimmt. Je größer die Mobilität eines Fahrers ist, desto mehr Knoten und Kanten sollte der entsprechende Graph enthalten. Für jeden Fahrer f werden Werte p_f und q_f für die Anzahl der Knoten und Kanten der jeweiligen Graphen bestimmt.

Da im Gegensatz zu den beiden vorherigen Perspektiven kein externer Vergleich möglich ist, wird zur Ermittlung der Kennwerte für jeden Fahrer f eine Min-Max-Normalisierung für die Werte der Stichprobe durchgeführt. Auf diese Weise wird ein Kennwert K_f des Fahrers relativ zu den anderen Fahrern bestimmt. Dafür wird das Maximum sowie das Minimum für die Mengen P und Q an Werten p_f und q_f für die Knoten und Kanten der Fahrer genutzt. Die Kennwerte K_p und K_q für die Knoten- und Kantenanzahl liegen im Intervall $[0,1]$. Die Berechnung der Kennwerte für jeden Fahrer (K_{p_f} und K_{q_f}) sind in den Gleichungen 5.8 und 5.9 dargestellt.

$$K_{p_f} = \frac{p_f - \min(P)}{\max(P) - \min(P)} \quad \forall \text{ Fahrer } f \quad (5.8)$$

$$K_{q_f} = \frac{q_f - \min(Q)}{\max(Q) - \min(Q)} \quad \forall \text{ Fahrer } f \quad (5.9)$$

Außerdem wird für jeden Fahrer der Mittelwert der Kantengewichte K_g ermittelt. Damit wird die Häufigkeit der Messung auf einer Kante in die Berechnung des Indikators Größe mit einbezogen. Die Formel für die Kennwertbestimmung der Kantengewichte ist in Gleichung 5.10 zu finden und nutzt die Menge aller Mittelwerte G .

$$K_{g_f} = \frac{g_f - \min(G)}{\max(G) - \min(G)} \quad \forall \text{ Fahrer } f \quad (5.10)$$

Für die Berechnung des Indikators $I_{Struktur}$ des Graphen werden zwei Maße verwendet. Das erste Maß ist die Anzahl der Kreise t im Graphen. Ein Kreis ist ein geschlossener Kantenzug. Die Anzahl soll die Variabilität in der Routenwahl abbilden. Eine Illustration ist in Abbildung 5.6 dargestellt.

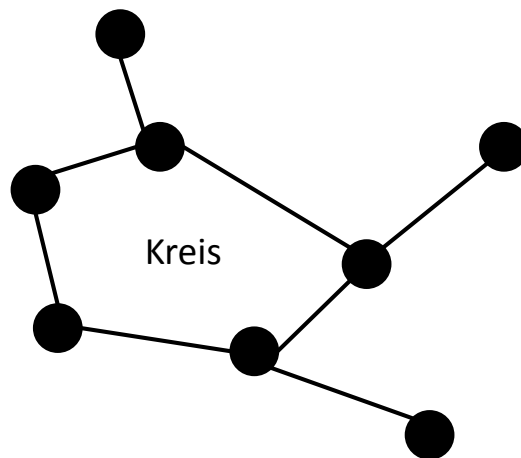


Abbildung 5.6: Illustration eines Kreises

Die Anzahl der Kreise t wird mit Hilfe der Eulerschen Polyederformel [93] bestimmt, siehe Gleichung 5.11.

$$t = q - p + 2 \quad (5.11)$$

Darauf baut die Kennwertbestimmung für die Anzahl der Kreise t_f für jeden Fahrer in Gleichung 5.12 auf. Auch hier wird wieder die Min-Max-Normalisierung genutzt, um den Kennwert in Bezug auf die Werte der Stichprobe T zu bestimmen.

$$K_{t_f} = \frac{t_f - \min(T)}{\max(T) - \min(T)} \quad \forall \text{ Fahrer } f \quad (5.12)$$

Der zweite und dritte Kennwert wird durch die Anzahl Knoten des Kerns eines Graphen bestimmt. Es wird zwischen dem *Kern* und dem *inneren Kern* unterschieden. Der Kern wird definiert als die größte Knotenuntermenge $p(\text{Kern})$ der Gesamtknotenmenge des Graphen, die durch Kanten mit einem Gewicht von mindestens zwei verbunden sind. Das heißt, dass eine oder mehrere inzidente Kanten mindestens zweimal durchfahren worden sind. Für die Knotenuntermenge des inneren Kerns $p(\text{inKern})$ muss der Kantengrad mindestens fünf sein. Somit sollen Kanten identifiziert werden, die mehr als dreimal in eine Richtung befahren wurden, wenn man gleiche Hin- und Rückwege zu einem Ziel annimmt. In Abbildung 5.7 ist ein Graph mit Kantengewichten dargestellt. Der Knoten des inneren Kerns (schwarz), sowie des Kerns (schwarz und grau) sind markiert.

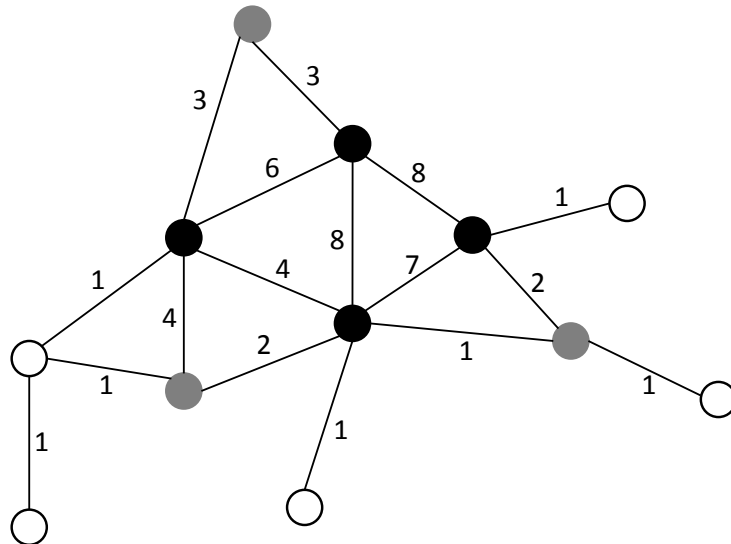


Abbildung 5.7: Illustration der Kerne (schwarz=innerer Kern, schwarz und grau=Kern)

Die Bestimmung des Kerns und damit die Identifizierung von Routinestrecken bzw. eines Bereichs mit Routineabschnitten ist komplexer, zeigt aber Strecken und Bereiche auf, auf denen Fahrer mehrfach gefahren sind. Damit ist dies ein wichtiger Parameter für die Struktur eines Graphen. So können Fahrer zwar viele Fahrten durchführen, dies aber häufig auf den gleichen Strecken, was sich an höheren Kantengewichten messen lässt. Sind diese Strecken zum Teil gleich oder kreuzen sich, sind auch die entsprechenden Kanten adjazent zueinander, und es bilden sich Knotenuntermengen der entsprechenden Knoten. Die Folge ist, dass die Größe des Graphen gering dafür aber die insgesamt gefahrene Strecke lang sein kann.

Die Kerne werden zunächst in Relation zum Gesamtgraphen gesetzt R_{Kern} und R_{inKern} , siehe Gleichungen 5.13 und 5.14. Damit soll ein Parameter dafür bestimmt werden, wie das Verhältnis zwischen dem Gebiet ist, das dem Fahrer vertraut ist und das er häufiger befährt und dem Gebiet mit Strecken, die während der Datenerhebung nur einmal befahren worden sind.

$$R_{Kern_f} = \frac{p_f(Kern)}{p_f} \forall \text{ Fahrer } f \quad (5.13)$$

$$R_{inKern_f} = \frac{p_f(inKern)}{p_f} \forall \text{ Fahrer } f \quad (5.14)$$

Die entsprechenden Kennwertbestimmungen für die Menge an Relativwerten aller Fahrer RK und $RinK$ mit Min-Max-Normierung folgen aus den Gleichungen 5.15 und 5.16:

$$K_{Kern_f} = \frac{R_{Kern_f} - \min(RK)}{\max(RK) - \min(RK)} \forall \text{ Fahrer } f \quad (5.15)$$

$$K_{inKern_f} = \frac{R_{inKern_f} - \min(RinK)}{\max(RinK) - \min(RinK)} \forall \text{ Fahrer } f \quad (5.16)$$

Nach der Bestimmung der Kennwerte erfolgt die Berechnung der Indikatoren. In den Indikator für die Größe $I_{Größe}$ gehen die Kennwerte für die Knotenanzahl, die Kantenanzahl sowie die Mittelwerte der Kantengewichte jeweils mit gleichem Gewicht ein, siehe Gleichung 5.17. Durch Normierung liegt auch dieser Indikator im Intervall $[0,1]$.

$$I_{Größe_f} = \frac{1}{3}(K_{p_f} + K_{q_f} + K_{g_f}) \forall \text{ Fahrer } f \quad (5.17)$$

In den Indikator $I_{Struktur}$ gehen die Kreisanzahl sowie die Kerne zusammen mit gleichem Gewicht ein. Damit haben die Kreise ein Gewicht von 0.5 sowie der Kern und der innere Kern ein Gewicht von jeweils 0.25. Die entsprechende Formel ist in Gleichung 5.18 dargestellt.

$$I_{Struktur_f} = \frac{1}{2}K_{t_f} + \frac{1}{4}(K_{Kern_f} + K_{inKern_f}) \forall \text{ Fahrer } f \quad (5.18)$$

Zur abschließenden Bestimmung des Navigationstyps werden die Fahrer in Abhängigkeit der beiden Indikatoren eingeordnet. Dafür wird ein Vergleich der Fahrer innerhalb der Stichprobe durchgeführt. Wir definieren vier verschiedene Navigationstypen in Abhängigkeit der Größe (Ausdehnung) und Struktur (Form), siehe Tabelle 5.2.

	Geringe Ausdehnung	Große Ausdehnung
Netzform	Lokal explorativer Navigationstyp	Explorativer Navigationstyp
Sternform	Lokal festgelegter Navigationstyp	Explorativ festgelegter Navigationstyp

Tabelle 5.2: Übersicht über die verschiedenen Navigationstypen

Der Graph von lokal festgelegten Typen ist tendenziell klein und besitzt nur wenig Kreise und kleine Kerne. Dementsprechend ist die Struktur sternförmig. Hingegen spannen die Graphen von explorativen Typen ein Netz auf und haben eine große Ausdehnung. Bei der Mischform lokal explorativ ist die Ausdehnung zwar gering, doch ist die Struktur netzartig. Im umgekehrten Fall haben Graphen von explorativ festgelegten Typen zwar eine große Ausdehnung, sind aber sternförmig. Eine Illustration der verschiedenen Typen ist in Abbildung 5.8 dargestellt.

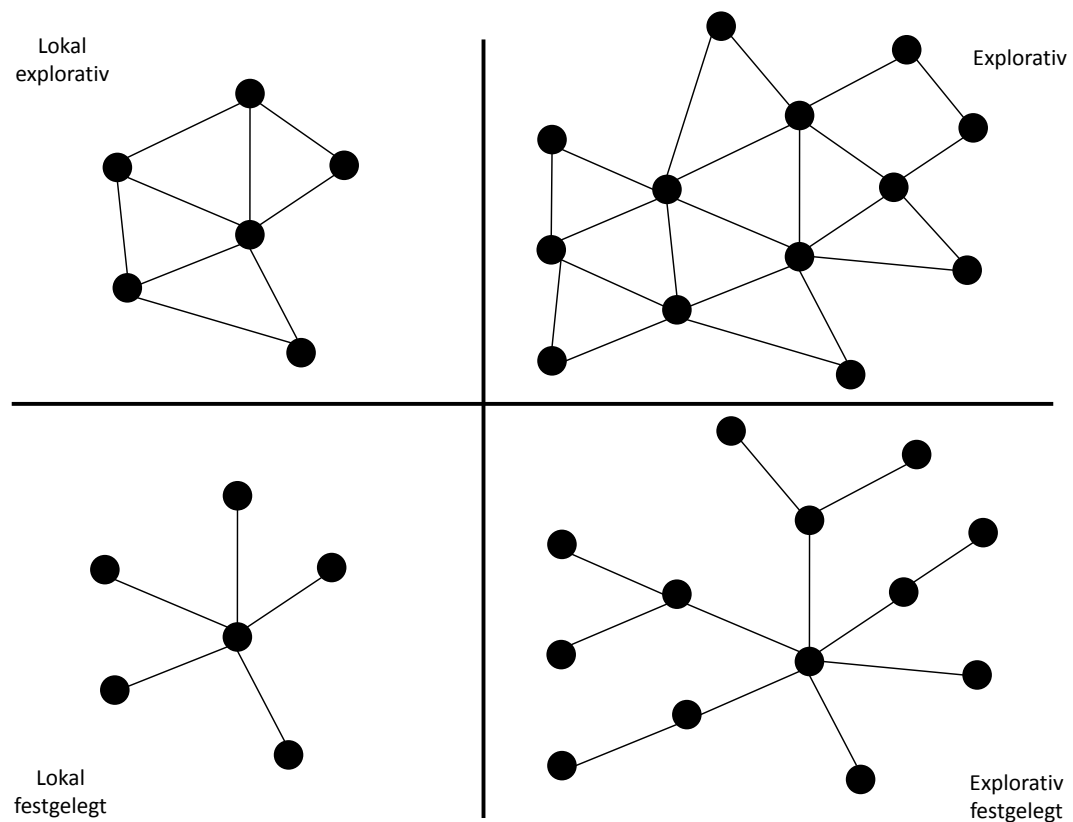


Abbildung 5.8: Illustration der verschiedenen Navigationstypen

Die Grenzen zur Einordnung werden durch die Mittelwerte über die Fahrer für die einzelnen Parameter bestimmt, für die eine Min-Max-Normalisierung der Stichprobe durchgeführt wird. Diese beschreibt die Einordnung eines Fahrers hinsichtlich des Parameters im Vergleich zu allen Fahrern.

Um ein Modell bezüglich der Dimension des Raumes zu erhalten, werden die beiden Indikatoren miteinander kombiniert. Zuerst werden die Indikatoren mit Hilfe der Min-Max-Normalisierung in Relation zu der Menge an Indikatoren $I_Größe$ und $I_Struktur$ der Gesamtstichprobe gesetzt, siehe Gleichung 5.19 und 5.20.

$$I_{GrößeMM_f} = \frac{I_{Größe_f} - \min(I_Größe)}{\max(I_Größe) - \min(I_Größe)} \quad \forall \text{ Fahrer } f \quad (5.19)$$

$$I_{StrukturMM_f} = \frac{I_{Struktur_i} - \min(I_Struktur)}{\max(I_Struktur) - \min(I_Struktur)} \quad \forall \text{ Fahrer } f \quad (5.20)$$

Zur Bestimmung eines entsprechenden Maßes M_a gehen beide normierten Indikatoren mit gleichem Gewicht ein, da beide den Graphen unabhängig voneinander beschreiben, siehe Gleichung 5.21. Das Ergebnis ist ein Wert, der die Abdeckung des Gebiets durch das erstellte Netzwerk des Fahrers auf Basis der Größe und der Struktur quantifizieren soll. Auch dieser Wert geht in die Korrelationsanalyse ein, mit dem Ziel ein zweidimensionales Modell zu erstellen.

$$M_{a_f} = \frac{1}{2}(I_{GrößeMM_f} + I_{StrukturMM_f}) \quad \forall \text{ Fahrer } f \quad (5.21)$$

5.3 Modellierung auf Basis der Ergebnisse der Perspektivenanalysen

5.3.1 Korrelationsanalyse

Nach der Analyse werden die Modelle der drei Perspektiven bzw. der zwei Dimensionen sowohl grafisch wie auch statistisch in Beziehung zueinander gesetzt. Für die grafische Analyse werden jeweils zwei Kennwerte der Fahrer untersucht. Für die statistische Analyse werden *Korrelationen* berechnet. Ziel beider Vorgehen ist die Ermittlung von Zusammenhängen zwischen den Modellen.

Für die statistische Analyse eignet sich die Schätzung der Korrelation mit dem Korrelationskoeffizient nach *Pearson* [12, 35]. Voraussetzung für die Bestimmung des *Pearson r* ist die *Normalverteilung* [12] der Werte. Ist diese nicht gegeben, muss auf analoge verteilungsfreie Verfahren zurückgegriffen werden.

Der Ablauf der Korrelationsanalyse, siehe Abbildung 5.9, ist an die Korrelationsbetrachtung aus Abbildung 3.6 angelehnt. In einem ersten Schritt werden die Maße M_v und M_d , die zur Dimension der Zeit sowie die Maße M_l und M_a , die zur Dimension des Raumes gehören, jeweils miteinander in Korrelation gesetzt. Anschließend erfolgt die Korrelationsanalyse für die Maße M_d und M_l , die beide der Routenperspektive zugeordnet sind. Ebenso wird eine dimensions- und perspektivenübergreifende Analyse durchgeführt. Die Ergebnisse zeigen das Potential und die Möglichkeiten einer zweidimensionalen Modellierung auf.

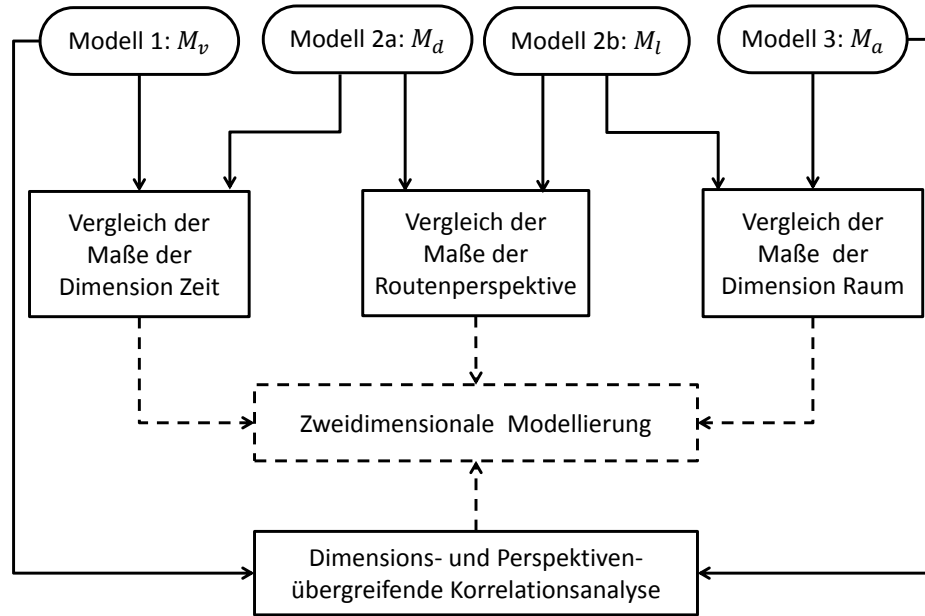


Abbildung 5.9: Ablauf der Korrelationsanalyse

5.3.2 Erstellung eines zweidimensionalen Modells

Bei der Modellierung wird jeder Fahrer einem *Routingtyp* zugeordnet. Der Ablauf der Modellierung ist in Abbildung 5.10 dargestellt. Die Modellierung baut auf den Ergebnissen auf, die die Analysen der verschiedenen Maße ergeben haben. Diese werden zunächst normiert. Die erhaltenen Maße M_{vMM} , M_{dMM} , M_{lMM} und M_{aMM} werden für die Erstellung eines zweidimensionalen Modells genutzt. Maße gleicher Dimension aus den vorherigen Analysen gehen gleichgewichtet ein. Das kombinierte Maß für die Dimension Zeit M_z ergibt sich aus den Werten M_{vMM} für die Geschwindigkeit der Kantenperspektive sowie der Dauer der Routenperspektive M_{dMM} . Die Berechnung für jeden Fahrer ist in Gleichung 5.22 zu finden.

$$M_{z_f} = \frac{1}{2}(M_{vMM_f} + M_{dMM_f}) \quad \forall \text{ Fahrer } f \quad (5.22)$$

Das Maß für den Raum M_r wird aus dem Maß der Länge M_{lMM} sowie dem Abdeckungsmaß M_{aMM} berechnet, siehe Gleichung 5.23. Das resultierende zweidimensionale Modell kann sowohl quantitativ wie auch deskriptiv beschrieben werden und bildet die Grundlage für die Personalisierung im nächsten Abschnitt.

$$M_{rf} = \frac{1}{2}(M_{lMM_f} + M_{aMM_f}) \quad \forall \text{ Fahrer } f \quad (5.23)$$

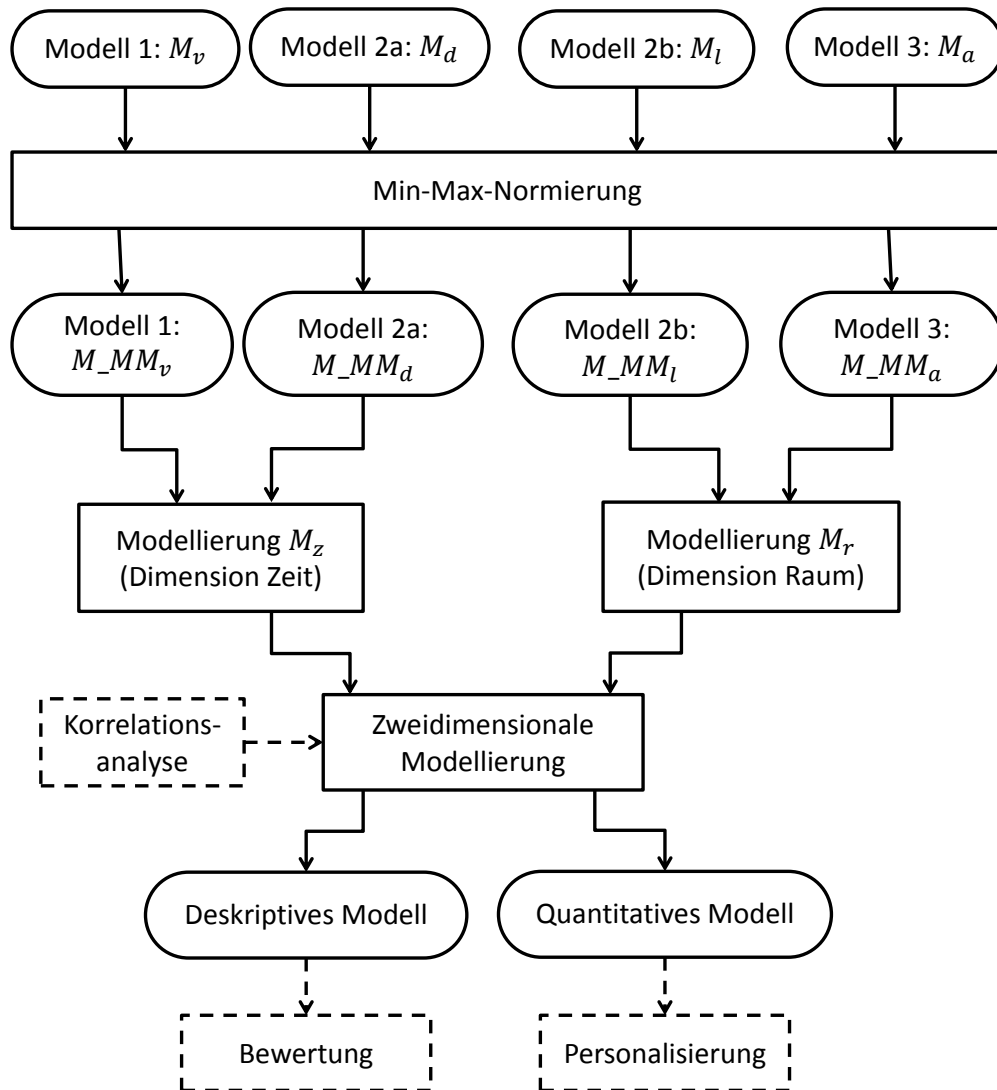


Abbildung 5.10: Ablauf der zweidimensionalen Modellierung

5.4 Durchführung der Personalisierung

Eine Übersicht über die einzelnen Personalisierungsansätze ist in Abbildung 5.11 dargestellt. Dabei fließen wie bei der Modellierung die Dimension der Zeit und

die Dimension des Raumes in unterschiedlicher Weise ein und werden miteinander kombiniert. Für jeden Fahrer wird jede Dimension personalisiert mit einem Anteil gewichtet. Das Vorgehen orientiert sich an der Vorstellung des Reglers, der in Abbildung 3.1 beschrieben wurde. Der Raumanteil kann dabei als fixer Bestandteil angesehen werden, da sich Längen nicht verändern können. Der Zeitanteil hingegen ist variabel. Die Variabilität ergibt sich aus den unterschiedlichen gemessenen Geschwindigkeiten der Fahrer.

Die Personalisierung der Kantengewichte der Geschwindigkeit, die Personalisierung der Zielfunktion sowie die Kombination der beiden Ansätze werden in den Abschnitten 5.4.1 bis 5.4.3 dargestellt.

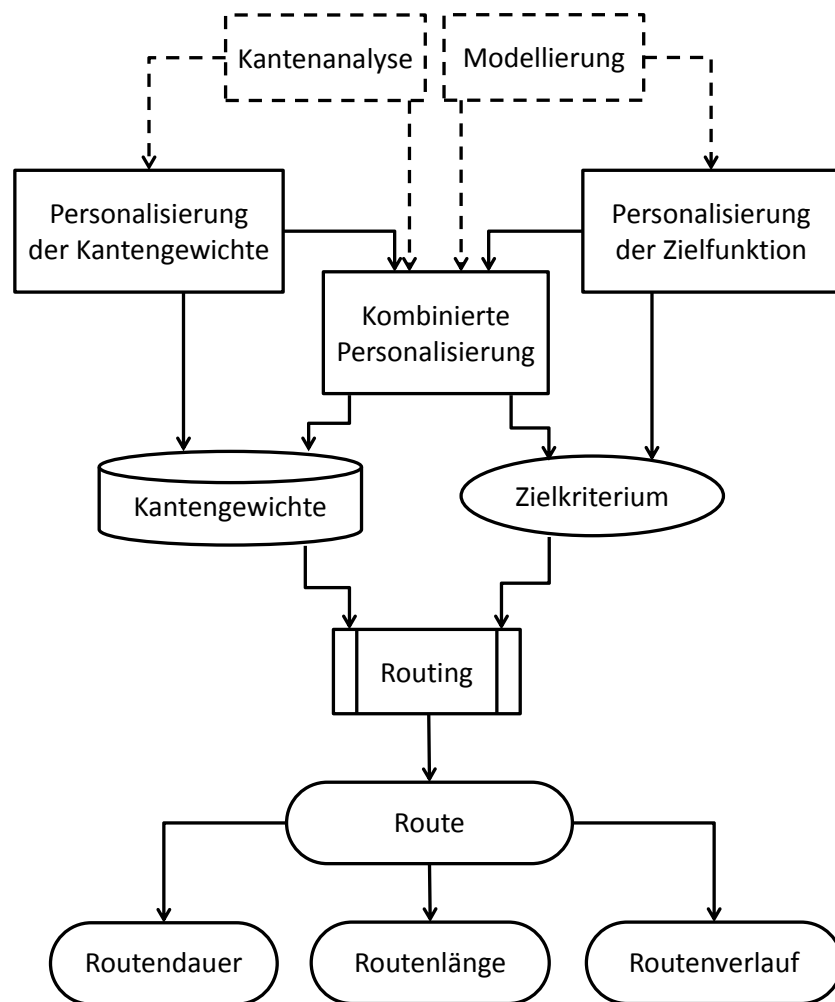


Abbildung 5.11: Visualisierung der unterschiedlichen Personalisierungsansätze

5.4.1 Ansatz der personalisierten Kantengewichte

Im ersten Schritt werden die Ergebnisse der Kantenanalyse genutzt, um die Kantengewichte zu personalisieren und an den Fahrer anzupassen. Hierfür wird, wie auch für die anderen Ansätze, eine *interne Routingreferenz* genutzt. Mit Hilfe der internen Routingreferenz werden Routen auf Basis einer Start-Zielkombination und einer digitalen Karte berechnet. Die Struktur der internen Routingreferenz ist in Abbildung 5.12 dargestellt. Es ist möglich, eine beliebige Zielfunktion anzugeben, wenn diese mit den Kantengewichten der Kanten der digitalen Karte abgebildet werden kann. Die digitale Karte ist eine selbst (intern) auf Basis des Realfahrdatensatzes erstellte Karte (siehe Abschnitt 6.3.4) und kann entsprechend mit beliebigen Kantengewichten versehen werden.

Es gibt ebenfalls die Möglichkeit die Kantengewichte, z.B. in Abhängigkeit der navigatorischen Präferenzen eines Fahrers, zu verändern und eine Personalisierung der Kantengewichte vorzunehmen. Die Eingaben nutzt die Routingkomponente, die auf einem KW-Verfahren basiert, für die Routenberechnung. Dafür wird das KW-Verfahren von Dijkstra [22] genutzt, um eine optimale Route zu berechnen. Damit stellt der zentrale Teil der Routingreferenz keine Black Box dar, sondern ermöglicht einen direkten Einfluss auf die digitale Karte als Grundlage oder die Berechnung selbst. Das Ergebnis der Routenberechnung ist ein Kantenzug, aus dem mit Hilfe der digitalen Karte Routencharakteristika abgeleitet werden können.

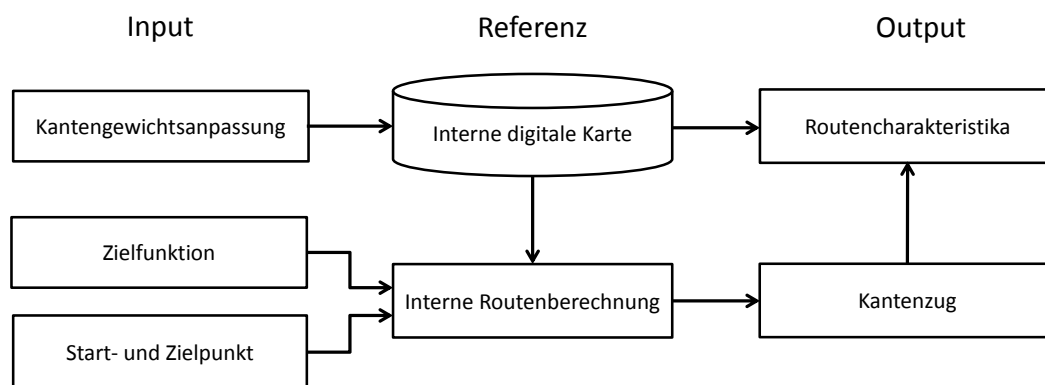


Abbildung 5.12: *Struktur der internen Routingreferenz*

Bei der Personalisierung der Kantengewichte wird ausschließlich die Dimension der Zeit berücksichtigt und entsprechend gewichtet. Die Individualkantengewichte, die für die Modellierung der Kantenperspektive berechnet wurden, werden für die

Routenberechnung benutzt und nutzen so das Potential des variablen Teils der Routenberechnung. Der Ablauf des Verfahrens ist in Abbildung 5.13 dargestellt.

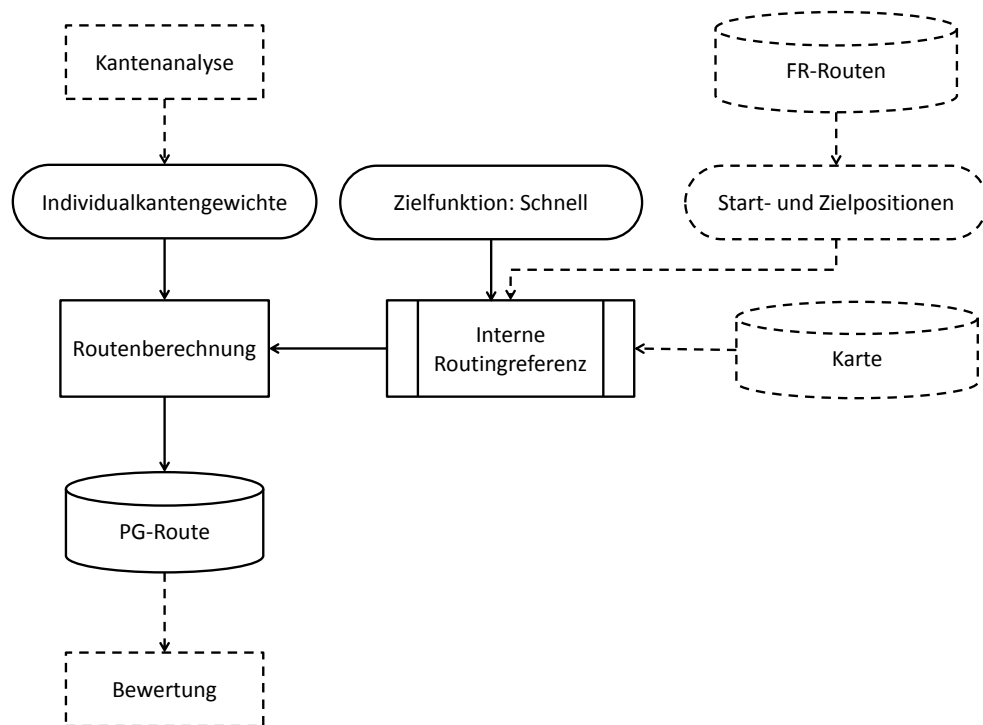


Abbildung 5.13: *Ablauf der Personalisierung der Kantengewichte*

Für die Routenberechnung werden die Karte sowie Start- und Zielpunkte für alle Fahrer-Realrouten bereitgestellt. Für die Start- und Zielpunkte der Systemrouten werden schnellste Routen basierend auf den Individualkantengewichten berechnet. Das Ergebnis der Personalisierung der Kantengewichte sind *PG-Routen*, die auf einer veränderten Datenbasis berechnet worden sind.

5.4.2 Ansatz der personalisierten Zielfunktion

Im Gegensatz zur Personalisierung der Kantengewichte wird bei der Personalisierung der Zielfunktion die Grundlage der Routenberechnung nicht verändert. Vielmehr wird die Zielfunktion für jeden Fahrer personalisiert und verschiedene Gewichte für den variablen und den fixen Anteil berechnet. Die Personalisierung der Zielfunktion basiert auf Erkenntnissen der Routenperspektive und berücksichtigt sowohl die Dimension der Zeit wie auch des Raumes. Bei der Analyse wurden zwei Kennwerte für die Maße M_z und M_r berechnet. Diese stellen somit eine Präferenz

für die eine oder andere Dimension dar. So gibt es z.B. Fahrer, die kürzere Strecken bevorzugen. Dies sollte sich auch in der Zielfunktion widerspiegeln, so dass nicht nur die benötigte Zeit berücksichtigt wird sondern auch die Länge der Route.

Das Verfahren ist in Abbildung 5.14 dargestellt. Für die Nutzung in der Zielfunktion werden aus den Kennwerten M_{z_f} und M_{r_f} der Fahrer Anteile ($M_{z_f}(Norm)$ und $M_{r_f}(Norm)$) berechnet, siehe Gleichung 5.24 und 5.25. Die personalisierte zweidimensionale Zielfunktion wird genutzt, um mit Hilfe der internen Routingreferenz *PZ-Routen* auf Basis der Standardkantengewichte zu berechnen.

$$M_{z_f}(Norm) = \frac{M_{z_f}}{M_{z_f} + M_{r_f}} \quad \forall \text{ Fahrer } f \quad (5.24)$$

$$M_{r_f}(Norm) = \frac{M_{r_f}}{M_{z_f} + M_{r_f}} \quad \forall \text{ Fahrer } f \quad (5.25)$$

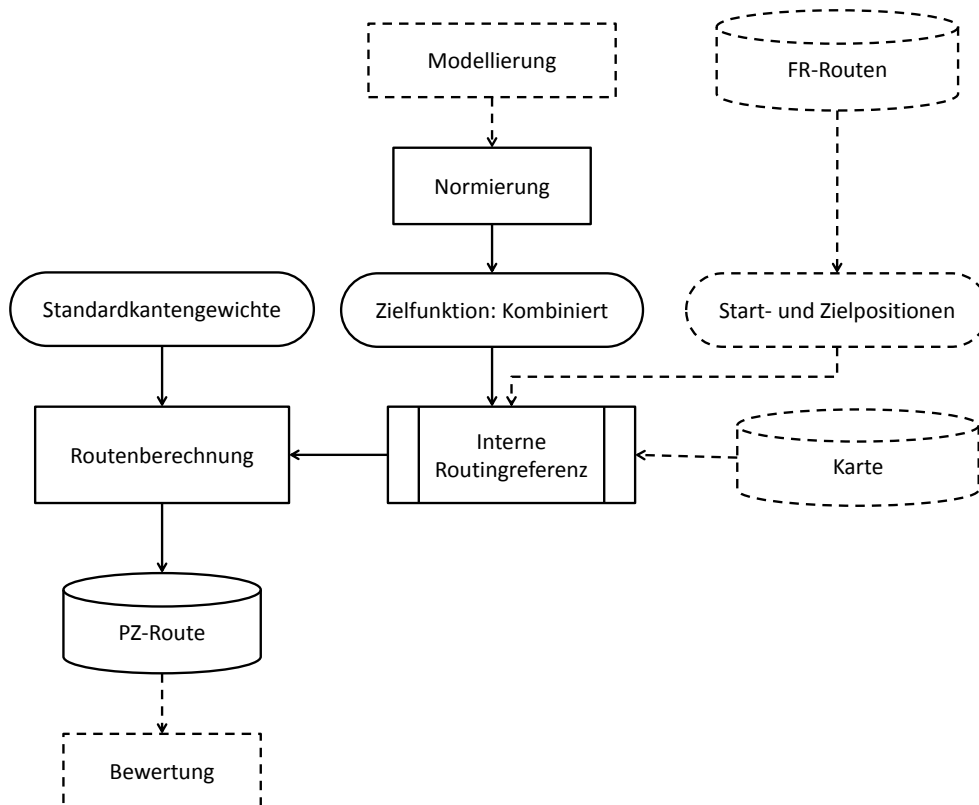


Abbildung 5.14: Ablauf der Personalisierung der Zielfunktion

5.4.3 Kombiniertes Personalisierungsansatz

Bei der kombinierten Personalisierung werden die beiden vorherigen Ansätze kombiniert und auf Basis einer personalisierten Datenbasis und für eine personalisierte

zweidimensionale Zielfunktion eine Route berechnet (*PK-Route*). Somit werden Gewichte für den variablen sowie den fixen Anteil bestimmt und für den variablen Teil für jeden Fahrer die Datenbasis für die Routenberechnung personalisiert. Die Personalisierung der Datenbasis basiert auf den Individualkantengewichten. Die Kennwerte werden erneut normiert und gehen in die Zielfunktion ein, für die die PK-Route berechnet wird. Eine Darstellung des Ansatzes ist in Abbildung 5.15 zu finden. Er verläuft analog zu den vorherigen Ansätzen und nutzt die interne Routingreferenz zur Berechnung der PK-Route.

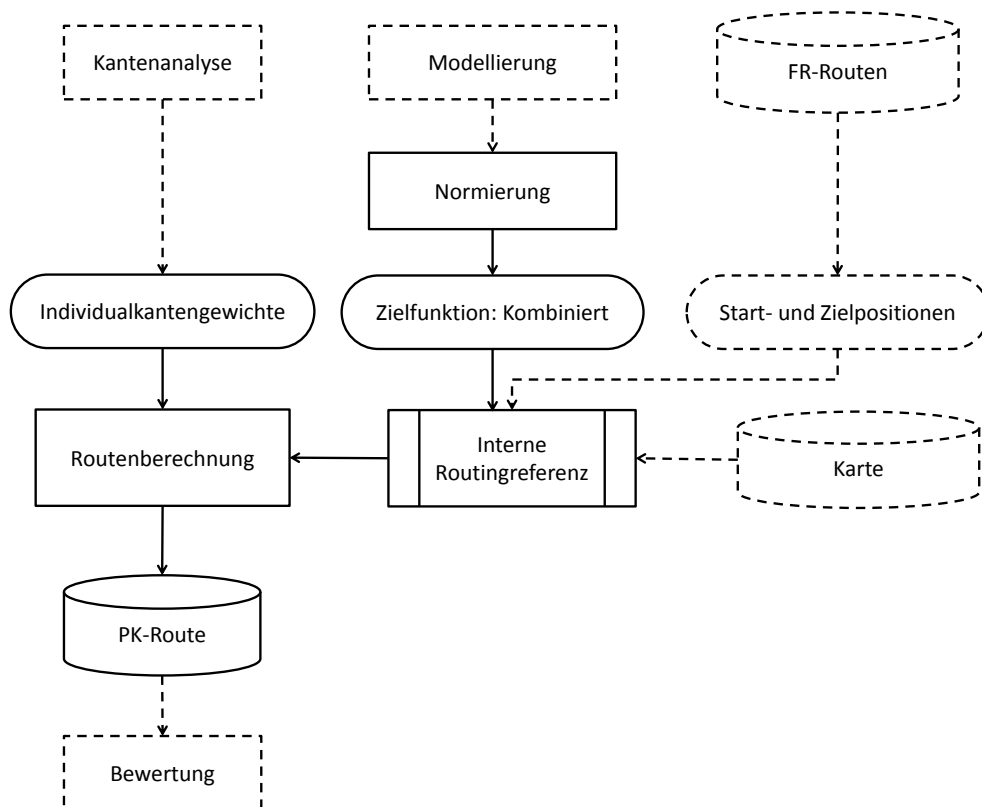


Abbildung 5.15: Ablauf der kombinierten Personalisierung

5.5 Bewertung der Personalisierung

Für die Bewertung der Personalisierung werden verschiedene Gütemaße genutzt, um die Veränderung durch die Personalisierung quantifizieren zu können. Grundlegendes Ziel der Personalisierung ist die Berechnung einer Route, die die ermittelten navigatorischen Präferenzen des Fahrers in höherem Maße berücksichtigt als eine Standardroute. Somit sollte eine solche Vorgehensweise eine Route zum Ergebnis haben, die der Fahrer entsprechend seiner Präferenzen mit größerer Wahrchein-

lichkeit wählen würde als die Standardroute. Die Quantifizierung kann durch Vergleiche von Routencharakteristika, wie die erreichten Zielfunktionswerte aber auch der Routen selbst erfolgen.

Für die Bewertung werden wie schon bei der Routenperspektive die FR-Routen genutzt, deren Länge mindestens 1 km beträgt. Bei kurzen Routen sind die Möglichkeiten der Variation gering. Es bietet sich nur eine Route an und auch eine Veränderung der Kantengewichte zieht nur marginale Änderungen nach sich. Für die Start- und Zielpositionen der FR-Routen können personalisierte Routen berechnet werden, die sich nicht nur in der Zeit sondern auch im Routenverlauf von der Standardroute unterscheiden können. Eine Übersicht über die Routen, die für die Bewertung genutzt werden, ist in Tabelle 5.3 zu finden.

Gemessene Routen	Berechnete Routen			
		Individualrouten		
FR-Routen	ST-Routen	PG-Routen	PZ-Routen	PK-Routen

Tabelle 5.3: Übersicht über die für die Bewertung genutzten Routen

Das erste Gütemaß $G_Z(vgl)$ der Zielkriterien wird auf Basis der Gütemaße $G_{Zd}(vgl)$ und $G_{Zl}(vgl)$ für die Routencharakteristika Dauer und Länge gebildet. Für das zweite Maß der Streckencharakteristika werden die Abweichungen der Anteile der Straßenklasse, z.B. der Anteil der Autobahnnutzung zwischen der FR-Route und den berechneten Routen bestimmt. Dabei wird zwischen einer kanten- und einer längenbasierten Untersuchung unterschieden (Gütemaße $G_{Sk}(vgl)$ und $G_{Sl}(vgl)$). Das dritte Gütemaß $G_{Uk}(vgl)$ vergleicht auf Kantenbasis die tatsächliche Übereinstimmung der berechneten Routen mit der FR-Route.

5.5.1 Gütemaß der Zielkriterien

Die Routendauer und die Routenlänge waren Gegenstand der Analysen zur Modellierung. Aus diesem Grund sollte eine Personalisierung zu einer Verbesserung im Bezug auf diese Kriterien führen. Eine Bewertung muss dabei auf Basis der Individualkantengewichte erfolgen, da diese auch die Grundlage für die Routenauswahlentscheidung der Fahrer bilden. Daher wird in einem ersten Schritt für die FR-Routen die theoretische Routendauer $d_r(obs_th)$ und Routenlänge $l_r(obs_th)$ berechnet, siehe Abbildung 5.16, links.

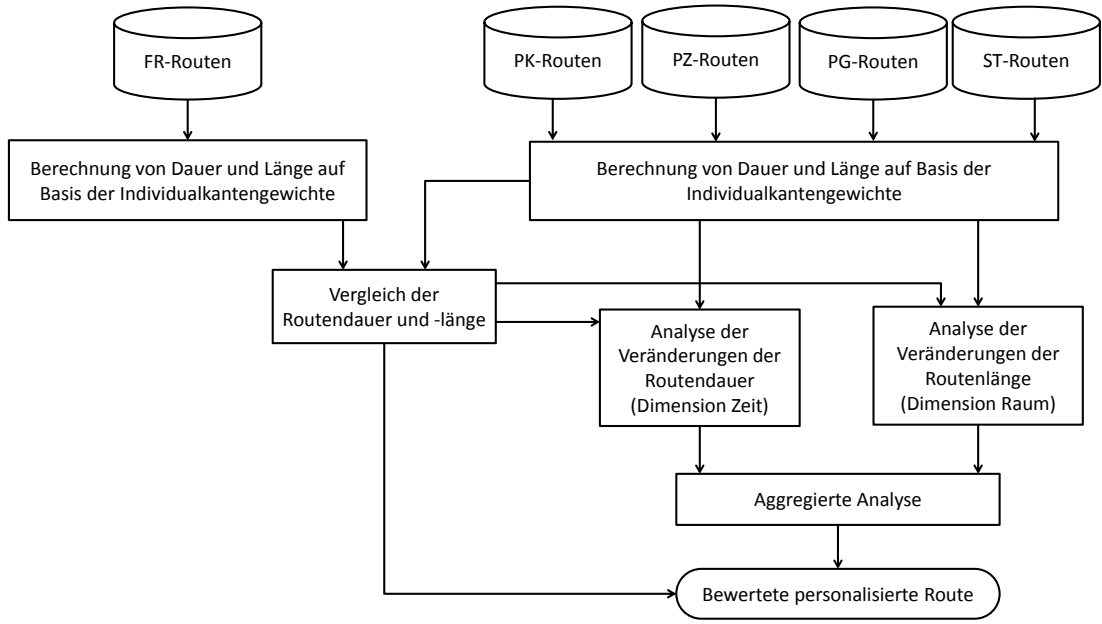


Abbildung 5.16: Vorgehen zur Bewertung auf Basis der Zielkriterien

Die theoretischen Werte der FR-Routen werden mit den entsprechenden Werten der berechneten Routen verglichen, um eine Einordnung der Werte zu erreichen. Der Vergleich erfolgt analog zu den vorherigen 0-Relativ-Vergleichen, siehe Formel 5.26 und 5.27.

$$g_Zd_{r,f} = \begin{cases} \frac{d_{r,f}(obs_th)}{d_r(calc)} - 1, & \text{if } d_{r,f}(obs_th) \leq d_r(calc) \\ 1 - \frac{1}{\frac{d_{r,f}(obs_th)}{d_r(calc)}}, & \text{if } d_{r,f}(obs_th) > d_r(calc) \end{cases} \quad \forall r \text{ und } f \quad (5.26)$$

$$g_Zl_{r,f} = \begin{cases} \frac{l_{r,f}(obs_th)}{l_r(calc)} - 1, & \text{if } l_{r,f}(obs_th) \leq l_r(calc) \\ 1 - \frac{1}{\frac{l_{r,f}(obs_th)}{l_r(calc)}}, & \text{if } l_{r,f}(obs_th) > l_r(calc) \end{cases} \quad \forall r \text{ und } f \quad (5.27)$$

Anschließend werden die Mittelwerte über die Menge aller Vergleichswerte G_Zd_f und G_Zl_f eines Fahrers gebildet, siehe Gleichung 5.28 und 5.29.

$$G_{Zd_f} = mean(G_Zd_f) \quad \forall \text{ Fahrer } f \quad (5.28)$$

$$G_{Zl_f} = mean(G_Zl_f) \quad \forall \text{ Fahrer } f \quad (5.29)$$

Diese können für eine Einordnung der berechneten Routen im Vergleich zu den FR-Routen genutzt werden. Dazu wird der Vergleich in Abhängigkeit der Routingtypen durchgeführt, um Unterschiede zwischen den Typen zu finden und damit die Modellierung zu validieren.

Nach der Einordnung wird die Veränderung durch die Personalisierung quantifiziert. Ausgehend von der PR-Route wird die Abweichung für beide Maße für die anderen drei berechneten Routen (PZ-Route, PG-Route, ST-Route) bestimmt. Zuerst werden die entsprechenden Maße $G_{Zd}(vgl)$ und $G_{Zl}(vgl)$ separat miteinander verglichen. Zur Bestimmung des Gütemaßes $G_Z(vgl)$ werden für jeden Fahrer die beiden Maße entsprechend der vorher bestimmten normierten Gewichtung $M_{z_f}(norm)$ und $M_{r_f}(norm)$ für die Dimensionen der Zeit und der Dimension des Raumes aggregiert.

Dafür wird erneut ein 0-Relativ-Vergleich für die einzelnen Routen der Fahrer durchgeführt, wobei die Fälle vertauscht werden, um den Einfluss des Informationsverlusts gegenüber der PR-Route korrekt darzustellen, siehe Gleichung 5.30 und 5.31.

$$g_Zd_{r,f}(vgl) = \begin{cases} \frac{d_{r,f}(PR)}{d_r(calc)} - 1, & \text{if } d_{r,f}(PR) > d_r(calc) \\ 1 - \frac{1}{\frac{d_{r,f}(PR)}{d_r(calc)}}, & \text{if } d_{r,f}(PR) \leq d_r(calc) \end{cases} \quad \forall r \text{ und } f \quad (5.30)$$

$$g_Zl_{r,f}(vgl) = \begin{cases} \frac{l_{r,f}(PR)}{l_r(calc)} - 1, & \text{if } l_{r,f}(PR) > l_r(calc) \\ 1 - \frac{1}{\frac{l_{r,f}(PR)}{l_r(calc)}}, & \text{if } l_{r,f}(PR) \leq l_r(calc) \end{cases} \quad \forall r \text{ und } f \quad (5.31)$$

Anschließend werden erneut die Mittelwerte über die Menge aller Vergleichswerte $G_Zd_f(vgl)$ und $G_Zl_f(vgl)$ eines Fahrers gebildet, siehe Gleichung 5.32 und 5.33.

$$G_{Zd_f}(vgl) = \text{mean}(G_Zd_f(vgl)) \quad \forall \text{ Fahrer } f \quad (5.32)$$

$$G_{Zl_f}(vgl) = \text{mean}(G_Zl_f(vgl)) \quad \forall \text{ Fahrer } f \quad (5.33)$$

Anschließend werden für jeden Fahrer die beiden Maße entsprechend der vorher bestimmten Gewichtung für die Dimensionen der Zeit und der Dimension des Raumes aggregiert, um das Gütemaß der Zielkriterien zu erhalten, siehe Gleichung 5.34.

$$G_{Z_f}(vgl) = M_{z_f}(norm) \cdot G_{Zd_f}(vgl) + M_{r_f}(norm) \cdot G_{Zl_f}(vgl) \quad \forall \text{ Fahrer } f \quad (5.34)$$

Zusätzlich wird noch eine Detailanalyse des Einflusses der einzelnen Perspektiven durchgeführt, siehe Abbildung 5.17. Dazu werden Informationen gegenüber der PK-Route weggelassen und die Auswirkungen für die so erhaltenen Routen analog zum vorherigen Vorgehen untersucht. Neben der PK-Route wird eine Route berechnet, für die die Information der Netzwerkperspektive für die Dimension des Raumes weggelassen wird. Dementsprechend werden nur Modell 1, 2a und 2b genutzt ($PK(1\&2a,2b)$ -Route). Eine weitere Route basiert nur auf den Erkenntnissen der Routenperspektive ($PK(2a,2b)$ -Route). Als drittes wird eine Route berechnet,

die für die Dimension des Raumes nur die Erkenntnisse der Netzwerkperspektive nutzt ($PK(1\&2a,3)$ -Route).

Auf die Darstellung der Gleichungen für den Vergleich wird an dieser Stelle verzichtet. Es wird die Veränderung im Vergleich zur PK-Route für alle drei Routen berechnet. Auf diese Weise soll noch einmal differenziert die Rolle der Dimension des Raumes und der entsprechenden Perspektiven untersucht werden.

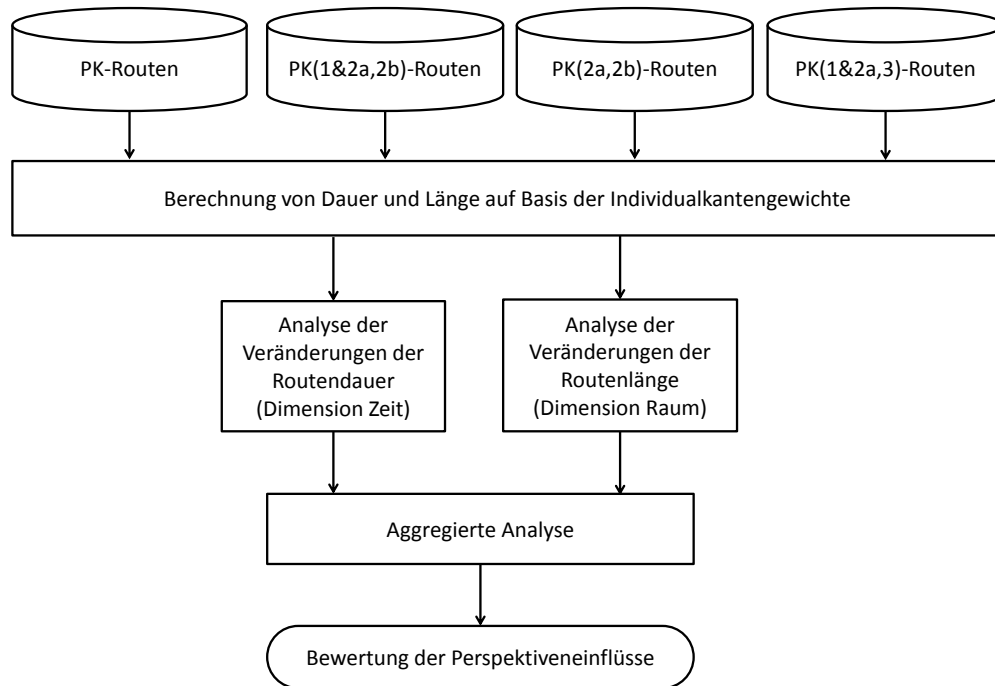


Abbildung 5.17: Vorgehen zur Analyse der Perspektivenmodelleinflüsse

5.5.2 Alternative Gütemaße

Gütemaß der Streckencharakteristika

Jede Route hat unterschiedliche Streckencharakteristika und einzelne Abschnitte können einer der drei Straßenklassen Autobahn (AB), Überland (UB) oder innerorts (IN) zugeordnet und Anteile für die Straßenklassen berechnet werden. Die Anteilsberechnung kann sowohl kanten- wie auch längenbasiert erfolgen. Die Bestimmung der Abweichung zwischen der FR-Route und einer berechneten Route wurde in Abbildung 3.3 illustriert. Das Vorgehen zur Bestimmung der Abweichungen der Straßenklassenanteile und somit zur Bewertung der Personalisierung ist in Abbildung 5.18 dargestellt.

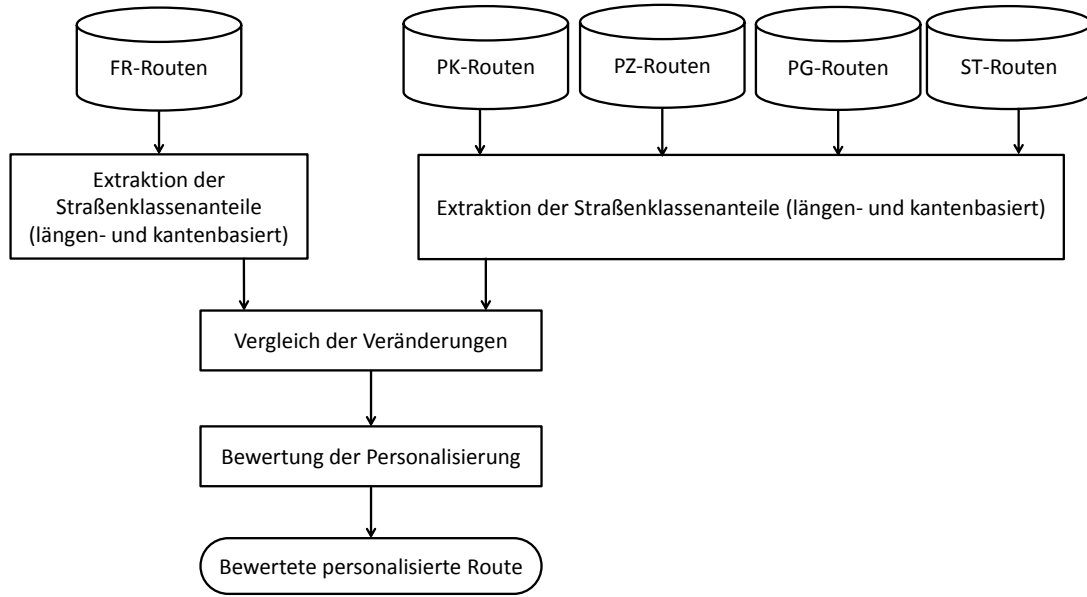


Abbildung 5.18: Vorgehen zur Bewertung auf Basis der Straßenklassenanteile

Die folgenden Formeln zeigen die kantenbasierte Berechnung für die Gütemaße G_{Sk} und $G_{Sk}(vgl)$. Längenbasiert erfolgt die Berechnung der Gütemaße G_{Sl} und $G_{Sl}(vgl)$ analog. In den Gleichungen 5.35, 5.36 und 5.37 sind die Berechnungen der Anteile $ANT_AB_{r,f}(obs)$, $ANT_UB_{r,f}(obs)$ und $ANT_IN_{r,f}(obs)$ der Straßenklassen Autobahn, Überland und innerorts einer gemessenen Route r zu finden.

$$ANT_AB_{r,f}(obs) = \frac{|Kanten_AB_{r,f}(obs)|}{|Kanten_{r,f}(obs)|} \quad \forall \text{ Routen } r \text{ und Fahrer } f \quad (5.35)$$

$$ANT_UB_{r,f}(obs) = \frac{|Kanten_UB_{r,f}(obs)|}{|Kanten_{r,f}(obs)|} \quad \forall \text{ Routen } r \text{ und Fahrer } f \quad (5.36)$$

$$ANT_IN_{r,f}(obs) = \frac{|Kanten_IN_{r,f}(obs)|}{|Kanten_{r,f}(obs)|} \quad \forall \text{ Routen } r \text{ und Fahrer } f \quad (5.37)$$

Die entsprechenden Anteile für die berechneten Routen $ANT_AB_{r,f}(calc)$, $ANT_UB_{r,f}(calc)$ und $ANT_IN_{r,f}(calc)$ werden nach den gleichen Formeln berechnet.

Im zweiten Schritt werden die Abweichungen der Straßenklassenanteile der berechneten von der gemessenen Route ermittelt. Hierfür wird die Absolutdifferenz genutzt und alle drei Klassen addiert, siehe Gleichung 5.38.

$$\begin{aligned} ANT_{r,f}(calc) = & abs(ANT_AB_{r,f}(calc) - ANT_AB_{r,f}(obs)) \\ & + abs(ANT_UB_{r,f}(calc) - ANT_UB_{r,f}(obs)) \\ & + abs(ANT_IN_{r,f}(calc) - ANT_IN_{r,f}(obs)) \end{aligned} \quad (5.38)$$

Um einen Kennwert G_{Sk_f} pro Fahrer für das entsprechende Gütemaß zu erhalten wird der Mittelwert über die Menge aller Abweichungen ANT_f eines Fahrers f gebildet, siehe Gleichung 5.39. Dieser Wert wird erneut zur Einordnung der berechneten Routen sowie zur Validierung der Modellierung genutzt.

$$G_{Sk_f} = \text{mean}(ANT_f) \quad \forall \text{ Fahrer } f \quad (5.39)$$

Für die Quantifizierung des Veränderung durch die Personalisierung wird analog zum Gütemaß der Zielkriterien die Veränderung gegenüber der PK-Route berechnet. Dafür wird ein 0-Relativ-Vergleich für die berechneten Abweichungen der Anteile $ANT_{r,f}$ durchgeführt, siehe Gleichung 5.40.

$$g-Sk_{r,f}(vgl) = \begin{cases} \frac{ANT_{r,f}(PK)}{ANT_{r,f}(calc)} - 1, & \text{if } ANT_{r,f}(PK) \leq ANT_{r,f}(calc) \\ 1 - \frac{1}{\frac{ANT_{r,f}(PK)}{ANT_{r,f}(calc)}}, & \text{if } ANT_{r,f}(PK) > ANT_{r,f}(calc) \end{cases} \quad (5.40)$$

Anschließend wird der Mittelwert über die Menge $G-Sk_f$ der erhaltenen Werte eines Fahrers berechnet, um so einen Wert für das Gütemaß der Streckencharakteristika zu bestimmen, siehe Gleichung 5.41. Somit wird die Veränderung der Abweichung der Anteile im Vergleich zur PK-Route bestimmt. Die Fallunterscheidung wurde so gewählt, dass eine größere Abweichung zu einem negativen Wert führt. Auf diese Weise zeigt die Richtung der Veränderung eine Verbesserung (positiv) oder eine Verschlechterung (negativ) an.

$$G_{Sk_f}(vgl) = \text{mean}(G-Sk_f) \quad \forall \text{ Fahrer } f \quad (5.41)$$

Übereinstimmung der Routen als Gütemaß

Abschließend wird das Gütemaß der Übereinstimmung G_{Uk} an Kanten der berechneten Routen mit der FR-Route beschrieben, siehe auch Abbildung 3.2. Daran schließt sich der Vergleich der Gütemaße der berechneten Route mit der PK-Route an ($G_{Uk}(vgl)$). Das Vorgehen für die Übereinstimmung verläuft analog zur Berechnung der Abweichung der Straßenklassenanteile, wobei dieses Mal ausschließlich eine Berechnung auf Basis der Kanten durchgeführt wird, siehe Abbildung 5.19.

Zuerst werden die Kanten der berechneten Route, die ebenfalls in der gemessenen Route vorkommen, bestimmt ($Kanten_Uk_{r,f}(calc)$). Anschließend wird die Anzahl dieser Kanten zu der Gesamtanzahl der Kanten der FR-Route in Relation gesetzt, siehe Gleichung 5.42.

$$ANT_Uk_{r,f}(calc) = \frac{|Kanten_Uk_{r,f}(calc)|}{|Kanten_{r,f}(obs)|} \quad \forall \text{ Routen } r \text{ und Fahrer } f \quad (5.42)$$

Für eine Einordnung der Übereinstimmungswerte und einer Bewertung der Modellierung wird der Mittelwert über die Menge an erhaltenen Werten ANT_Uk_f eines Fahrers berechnet, siehe Gleichung 5.43.

$$G_{Uk_f} = \text{mean}(ANT_Uk_f) \quad \forall \text{ Fahrer } f \quad (5.43)$$

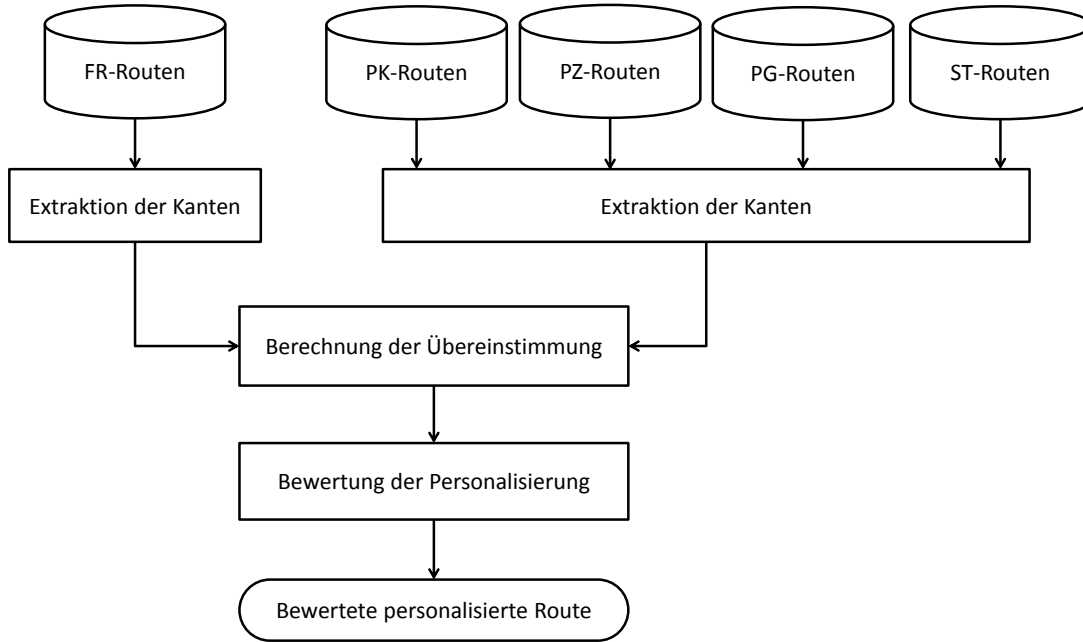


Abbildung 5.19: Vorgehen zur Bewertung auf Basis der Übereinstimmung

Die PK-Route dient wieder als Vergleich zur Quantifizierung des Nutzens. Es wird ein 0-Relativ-Vergleich für alle berechneten Routen über alle Fahrer durchgeführt, siehe Gleichung 5.44.

$$g_Uk_{r,f} = \begin{cases} \frac{ANT_Uk(calc)_{r,f}}{ANT_Uk(PK)_{r,f}} - 1, & \text{if } ANT_Uk(calc)_{r,f} \leq ANT_Uk(PK)_{r,f} \\ 1 - \frac{1}{\frac{ANT_Uk(calc)_{r,f}}{ANT_Uk(PK)_{r,f}}}, & \text{if } ANT_Uk(calc)_{r,f} > ANT_Uk(PK)_{r,f} \end{cases} \quad (5.44)$$

Die Vergleichswerte werden erneut so gewählt, dass eine Verschlechterung zu einem negativen Vergleichswert führt. Im Fall der Übereinstimmung bedeutet dies, dass der Wert der Übereinstimmung kleiner ist, im Gegensatz zur Abweichung der Straßenklassenanteile.

Anschließend wird der Mittelwert über die Menge G_Uk_f der Relativvergleichswerte gebildet, siehe Gleichung 5.45. Somit wird ein Vergleich der berechneten Routen ermöglicht und zeigt die Veränderung durch die Personalisierung auf.

$$G_{Uk_f}(vgl) = \text{mean}(G_Uk_f) \quad \forall \text{ Fahrer } f \quad (5.45)$$

6 Datenaufbereitung

In diesem Kapitel werden die Datenaufbereitungsschritte dargestellt, die für die Daten des euroFOT-Projekts notwendig waren. Es werden drei Datensätze für die Analyse der verschiedenen Perspektiven erstellt. Darüber hinaus wird eine Karte für die in Abschnitt 5.4.1 beschriebene interne Routingreferenz generiert, die für die Personalisierung genutzt wird. Die Aufbereitung umfasst die Prozessschritte Datenauswahl, Datenbereinigung und Datenreduktion, siehe Abbildung 6.1.

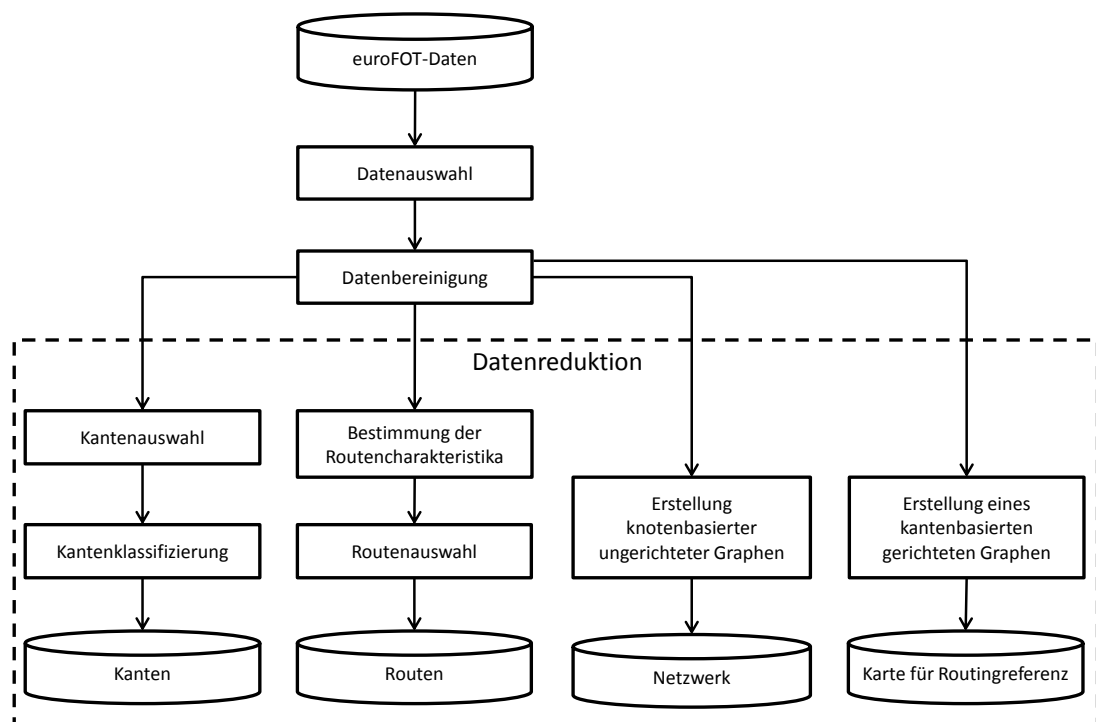


Abbildung 6.1: *Ablauf der Datenaufbereitung*

Die Datenauswahl erfolgt anhand der in Kapitel 3 dargestellten Anforderungen. Bei der anschließenden Bereinigung wird besonderen Wert auf die Ortung und Verbesserung des Matchings gelegt. Die Fahrdynamikdaten wurden bereits im Projekt euroFOT bereinigt. Ortung und Matching bestimmen, welche Messpunkte welcher Kante mit den entsprechenden Kanteninformationen der digitalen Karte zugeord-

net werden. Damit haben diese Bereinigungsschritte einen fundamentalen Einfluss auf die folgenden Reduktionsschritte.

Für die Reduktion auf einen Datensatz der Kanten werden diese zunächst in jeder Route identifiziert, für jeden Fahrer zusammengefasst und abschließend nach bestimmten Kriterien ausgewählt. Für die Routen, als geschlossener Kantenzug, werden zunächst Start- und Zielpunkt identifiziert. Dieses ermöglicht die Bestimmung der Routencharakteristika wie Routendauer und Routenlänge ebenso wie die Ermittlung des Routenverlaufs. In Abhängigkeit der Charakteristika werden geeignete Routen ausgewählt und bilden den Datensatz für die Routenanalyse aber auch als Input für die Berechnung der Routen mit Hilfe der Routingreferenz. Für die Bestimmung des Netzwerks werden zusätzlich die Zusammenhänge der Kanten bzw. Knoten bestimmt, um für jeden Fahrer ein Netzwerk als die Gesamtheit der gemessenen Routen zu konstruieren. Ebenso wird das gesamte Netzwerk der Kanten aller Fahrer zur Erstellung der Karte für die Routingreferenz genutzt.

Für die Aufbereitung der Daten des Projektes euroFOT wurde hierfür wie auch für die Auswertungen die Analysesoftware Matlab [3] verwendet. Matlab eignet sich für die Verarbeitung und Analyse großer Datenmengen, die in Form einer Matrix gespeichert sind. Für die Analyse ist es möglich, Programme in einer Programmiersprache zu schreiben, die Komponenten der objektorientierten Programmierung beinhaltet.

Jede Fahrt wurde im Projekt als eigene Matlab-Datei gespeichert und lagen als mat-File vor. Dementsprechend arbeiten die in dieser Arbeit programmierten Skripte für die Analysen mit mat-Files. Die Syntax lässt sich allerdings leicht auf andere Dateiformate übertragen.

Die statistischen Auswertungen sowie die Grafiken wurden größtenteils mit der Statistiksoftware Statistica [98] durchgeführt bzw. erstellt. Statistica bietet eine Vielzahl an deskriptiven und statistischen Analysewerkzeugen. Dazu bietet Statistica viele Möglichkeiten für die Erstellung und Anpassung von Grafiken auf Basis umfangreicher Daten in Tabellenform an.

6.1 Datenauswahl

Von den 115 in euroFOT erhobenen Fahrern werden die Fahrer ausgewählt, für die während des dreimonatigen Versuchszeitraums erfolgreich Daten erhoben und aufbereitet wurden. Diese Fahrer haben das Fahrzeug über einen ähnlich langen Zeitraum genutzt und können in die Analysen mit einbezogen werden, um ver-

gleichbare Ergebnisse zu erzielen. Für die Daten dieser Fahrer werden mehrere Datenauswahlschritte durchgeführt, siehe Abbildung 6.2.

Zunächst werden alle Fahrten ausgewählt, in denen sich das Fahrzeug bewegt hat und damit eine potentielle Route aufweisen. Da die Messtechnik mit dem Öffnen der Tür hochgefahren worden ist und nach dem Ausschalten des Motors noch ein gewisser Nachlauf bis zum Ende der Datenaufnahme eingeplant wurde, werden im nächsten Schritt Anfang und Ende der Route bestimmt, so dass die ausgewählten Abschnitte die Route des Fahrers beschreiben. Dafür wird die gemessene Geschwindigkeit genutzt. Die Fahrt beginnt, wenn sich das Fahrzeug zum ersten Mal bewegt und endet beim letzten Stopp des Fahrzeugs während der Aufzeichnung. Diese Zeitpunkte werden mit Hilfe der aufgezeichneten Geschwindigkeit bestimmt.

Damit wird für viele Fahrten am Anfang und Ende ein Zeitraum abgeschnitten, in der sich die Position nicht ändert und der für die Analyse der Routen irrelevant ist. Ein zu langer Vorlauf (z.B. durch Packen des Autos) könnte die Ergebnisse sogar verzerren und zu einer deutlich längeren gemessenen Routenzeit führen.

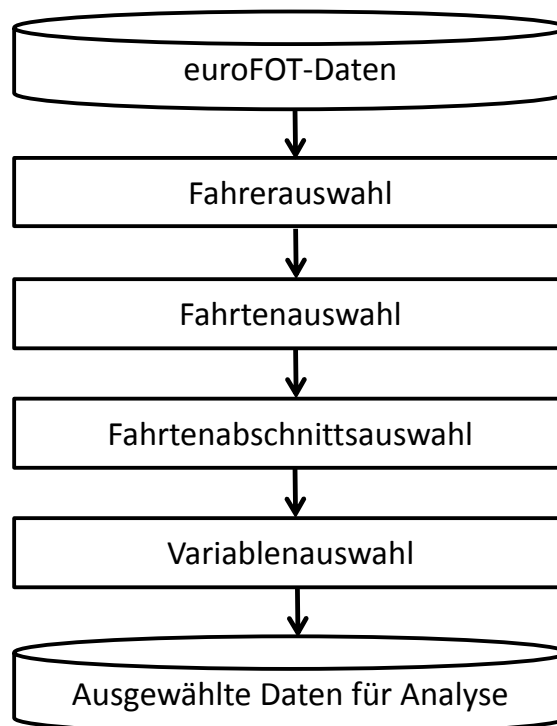


Abbildung 6.2: Ablauf der Datenauswahl

Im euroFOT-Projekt wurde eine Vielzahl von Daten aufgenommen. Aufgrund der Fokussierung auf Fahrerassistenzsysteme wurden viele Daten zur Beschreibung und Bewertung der Systeme erhoben. Aus den so zu jedem Messzeitpunkt erhaltenen

Signalen bzw. Daten werden die für diese Arbeit relevanten ausgewählt. Diese umfassen mehrere Blöcke, siehe Tabelle 6.1.

Datenblock	Variable	Erklärung
Fahrtdaten	VPNr	Nummer des Fahrers der Fahrt
	FahrtNr	Nummer der Fahrt eines Fahrers
	Bekanntheit	Angabe, ob eine Route bekannt oder unbekannt ist
Fahrtverlaufsdaten	Unix-Zeit	Absolute Zeit
	time	Zeitverlauf während der Fahrt
	gef_km	Zurückgelegte Strecke zwischen zwei Messpunkten
	gef_km_kum	Kumuliert zurückgelegte Strecke einer Fahrt
	Status Fahrzeug	Bewegungszustand des Fahrzeugs
	Status Motor	Aktivitätszustand des Motors
	Straßentyp	Aktuelle Straßenklasse
	v	Aktuelle Geschwindigkeit
Navigation	Status Navi	Aktivitätszustand des Navigationssystems
GPS-Daten	GPS_lat	Latitudinale Position des Fahrzeugs
	GPS_long	Longitudinale Position des Fahrzeugs
	GPS_heading	Ausrichtung des Fahrzeugs
	GPS_year	Jahr laut GPS-Signal
	GPS_month	Monat laut GPS-Signal
	GPS_day	Tag laut GPS-Signal
	GPS_hour	Stunde laut GPS-Signal
	GPS_minute	Minute laut GPS-Signal
	GPS_sec	Sekunde laut GPS-Signal
Kartendaten	MM.link_ID	ID der aktuellen Kante
	MM_lat	Latitudinale Position auf Kante
	MM_long	Longitudinale Position auf Kante
	MM.speedLimit	Geschwindigkeitsbegrenzung der aktuellen Kante
	MM.averageSpeed	Angenommene durchschnittliche Geschwindigkeit der aktuellen Kante
	MM.fromRefM	Durchfahrene Richtung der aktuellen Kante
	MM.allowedDirection	Erlaubte Richtungen der aktuellen Kante

Tabelle 6.1: *Ausgewählte Variablen der Datenbasis*

Statische Fahrtdaten geben Auskunft über Fahrer und Fahrzeug-ID, den Versuchsmonat, aber auch Angaben des Fahrers zur Bekanntheit der Strecke. Diese Informationen bleiben über eine Fahrt entlang einer Route konstant. Desweiteren sind Fahrtverlaufsdaten, insbesondere die Geschwindigkeit, ausgewählt worden. Als ein-

zuges relevantes Fahrzeugsystem werden Signale über den Systemzustand des Navigationssystems genutzt. Darüber hinaus sind die GPS-Daten, wie z.B. Position oder Uhrzeit, vorhanden. Einige grundlegende, in euroFOT ermittelte, Kartendaten werden ebenfalls genutzt.

Eine deskriptive Beschreibung der ausgewählten Daten ist in Tabelle 6.2 aufgeführt. Der Datensatz beinhaltet 106 von 115 Fahrern aus euroFOT. Dazu wurden einige Fahrten mit invaliden Werten ausgeschlossen. Dementsprechend haben sich die Fahrten um 3.3% auf 38 377 Fahrten verringert. Gleiches gilt für die Kilometer (Verringerung um 2.8%) und insbesondere für die Fahrtstunden (Verringerung um 3.7%). Mit über 100 Fahrern und mehr als 980 000 gefahrenen Kilometern bildet die Datenbasis allerdings weiterhin eine umfangreiche Datenbasis für die Analysen des Routenauswahlverhaltens der Fahrer.

	Fahrzeugtyp 1	Fahrzeugtyp 2	Gesamt
Anzahl Probanden	59	47	106
Anzahl Fahrten	23 537	14 840	38 377
Insgesamt gefahrene Kilometer	614 183	370 362	984 545
Aufgezeichnete Fahrtstunden	8 810	5 756	14 566

Tabelle 6.2: *Kennzahlen der Datenbasis nach der Datenauswahl*

6.2 Datenbereinigung

Die Bereinigung der erhobenen Positionsdaten erfolgt in vier Schritten für jede Fahrt, siehe Abbildung 6.3. Ziel der Bereinigung ist es, ein korrektes und plausibles Matching der Messpunkte zu Kanten der digitalen Karte zu erreichen. Damit einher geht die korrekte Zuordnung zur digitalen Karte.

Das entsprechende Verfahren umfasst die Identifizierung und Bereinigung von Lücken mit invaliden Positions- sowie Kartendaten, die Bearbeitung von Sprüngen der Position zu Beginn oder während der Fahrt und abschließend eine Bereinigung des in euroFOT durchgeführten Matchings.

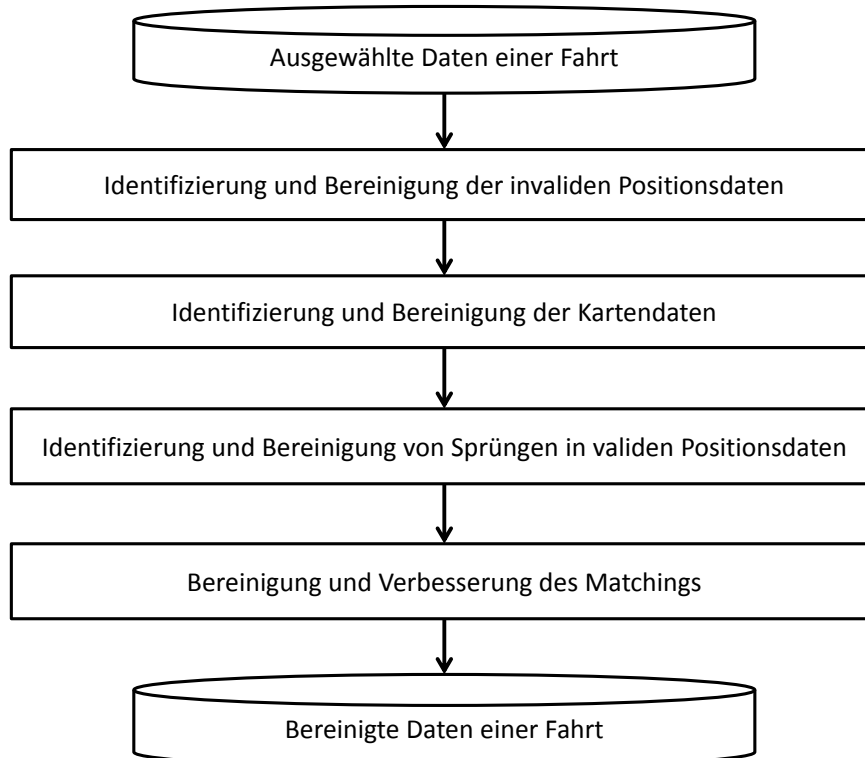


Abbildung 6.3: *Ablauf der Datenbereinigung für eine Fahrt*

6.2.1 Invalide Positionsdaten

Bei nicht möglichen Ortungen (z.B. Tunnel) springt die Position auf einen Defaultwert. Auch kann es zu kurzzeitigen Aussetzern kommen, die in euroFOT nicht bereinigt wurden. Solche Lücken in den Daten werden identifiziert, indem Beginn und Ende markiert werden, an dem die Position unplausibel wird bzw. letztmalig ist. Das anschließende Verfahren zur Bereinigung der Lücken ist in Abbildung 6.4 dargestellt.

Insgesamt wurden knapp 8500 Lücken identifiziert. Dabei ist darauf hinzuweisen, dass in diesem Schritt ausschließlich Lücken bereinigt werden, so dass bereits eine valide Position vorliegen muss. Abschnitte mit invaliden Werten zu Beginn werden in einem späteren Schritt bereinigt. Nach der Identifizierung werden die Lücken in den Positionsdaten für die Bereinigung charakterisiert. Als Kriterien werden die Dauer und Länge der Lücke, die Geschwindigkeit sowie die Veränderung der Position während der Lücke genutzt. Ungefähr 6600 der Lücken sind maximal eine Sekunde lang. Diese werden als Datenaussetzer klassifiziert und durch Fortschreiben der vorherigen Informationen über Position und Kartendaten geschlossen.

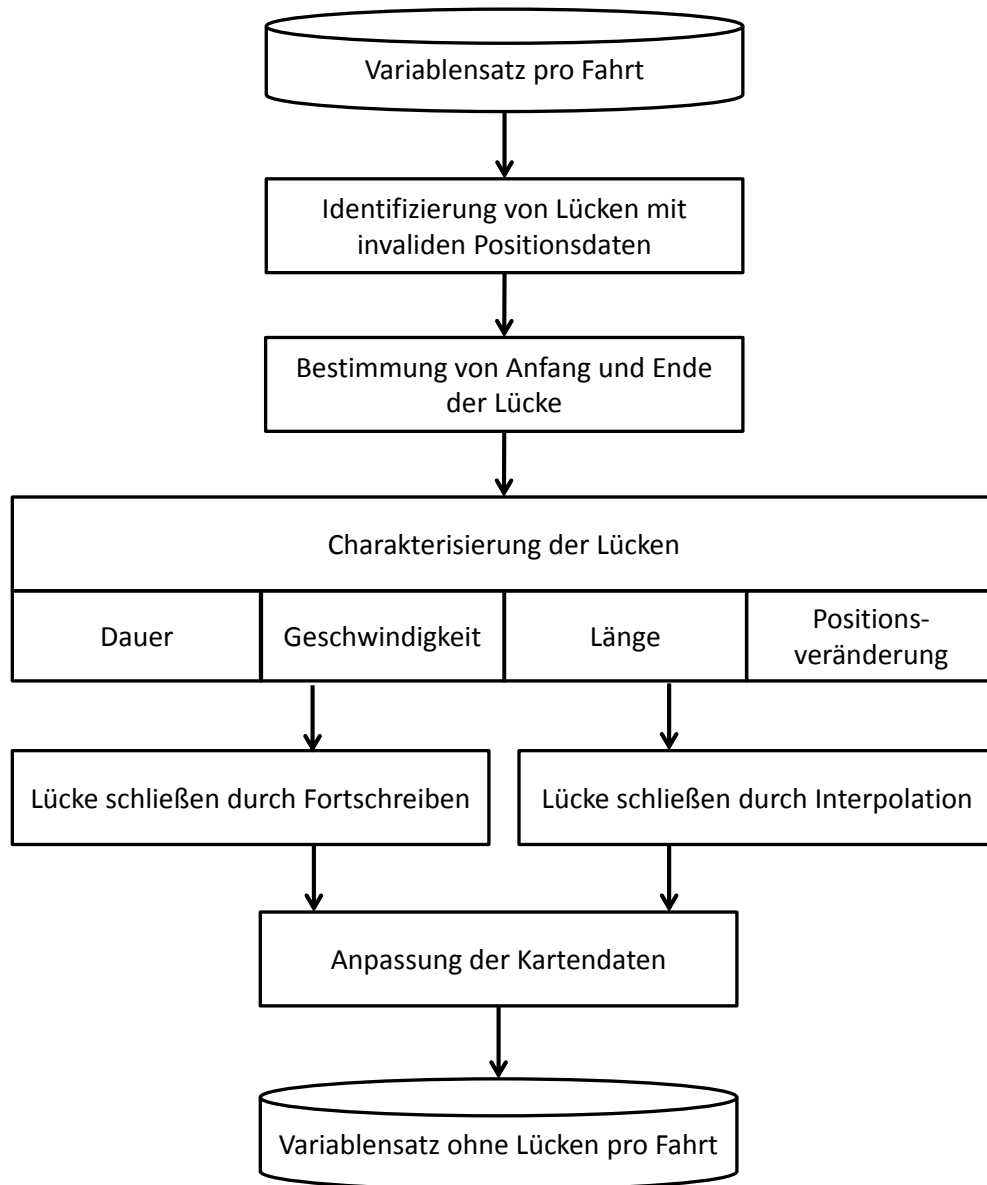


Abbildung 6.4: Vorgehen zur Bereinigung der Positionsdaten

Ist die Lücke länger, so wird zunächst die Geschwindigkeit während der Lücke geprüft. Ist diese Null, so wird die Position fortgeschrieben. Sind dazu die Kanten vor und nach der Lücke gleich, werden auch die Kartendaten fortgeschrieben. Unterscheiden sich die Kanten, so wird die Lücke zu gleichen Teilen zwischen den Kanten aufgeteilt, da der Sprung nicht eindeutig aufgeklärt werden kann. Dieser Fall tritt 150 Mal ein, gleiche Kanten 650 Mal.

Hat sich das Fahrzeug bewegt, wird die Veränderung der Position geprüft. Bei einer geringen Positionsveränderung von maximal 100 Metern wird die Lücke durch Interpolation geschlossen. Die entsprechenden Kanteninformationen werden vom

Anfang und vom Ende aus zu gleichen Teilen fortgeschrieben, wenn die Kanten unterschiedlich sind. Bei gleichen Kanten werden die Informationen fortgeschrieben. Dieser Fall kommt 600 Mal vor, verschiedene Kanten 370 Mal.

Bei einer größeren Positionsveränderung wird geprüft, ob sich vor und nach der Lücke die gleiche Kante befindet. Für diese 15 Fälle werden die Positionen interpoliert und die Karteninformationen fortgeschrieben. In 31 Fällen tritt die Lücke am Ende der Fahrt auf. Dies kann z.B. beim Befahren eines Parkhauses passieren. In diesen Fällen wird die Datenaufzeichnung entsprechend gekürzt.

Für die restlichen 56 Fälle werden zunächst Anfang und Ende der Lücke geprüft und unter Umständen angepasst, um sicherzustellen, dass die entsprechenden Kanten valide sind. Danach wird geprüft, ob die Positionen interpoliert und die Lücke auf die beiden Kanten aufgeteilt werden kann. Entscheidendes Kriterium ist die Länge der gefahrenen Strecke während der Lücke sowie auf den Kanten davor und danach. Stimmen diese mit den theoretischen Längen der Kanten überein, können diese zugeordnet werden. Auf diese Weise können für alle Fahrten bis auf drei (diese werden ausgeschlossen) die Positionen und entsprechend die Kartendaten bereinigt werden.

6.2.2 Kartendaten

Im nächsten Schritt werden die Kartendaten der Messpunkte bereinigt, die zwar eine valide Position besitzen aber beim Map Matching in euroFOT Kriterien für ein valides Matching nicht erfüllt haben, siehe Abbildung 6.5. Dies kann vorkommen, da ein rudimentärer Ansatz verwendet wurde, der die Historie der Position nicht oder nur in geringem Maßen nutzt und somit ungenaue Ortungen nicht kompensieren kann. Damit konnten keine Kartendaten extrahiert werden, die damit invalide sind.

Insgesamt werden ca. 5 200 Lücken identifiziert. Hierbei ist allerdings anzumerken, dass häufig mehrere Lücken einem Ereignis zugeordnet werden können und meist sehr kurz sind. Wenn ein Abschnitt mit invaliden Kartendaten identifiziert wurde, wird zunächst nach einem stabilen Matching von mehr als fünf Sekunden vor und nach dem Abschnitt gesucht. Zwei stabile Matchings davor und danach werden in 2800 Fällen gefunden, nur ein stabiles Matching danach in 2600 Fällen. Dies sind häufig Lücken am Anfang einer Fahrt. Der Fall, dass nur vor einer Lücke ein stabiles Matching gefunden werden kann, tritt 670 Mal auf.

Als nächstes wird die *Aufnahmefähigkeit* der gematchten Kanten geprüft. Hierzu wird die befahrene Strecke auf den beiden Kanten und der Lücke mit den theoretischen Längen der beiden Kanten verglichen. Sind diese länger als die gemessene Gesamtstrecke, so wird die Lücke entsprechend aufgeteilt. Das Verfahren für ein stabiles Matching verläuft analog.

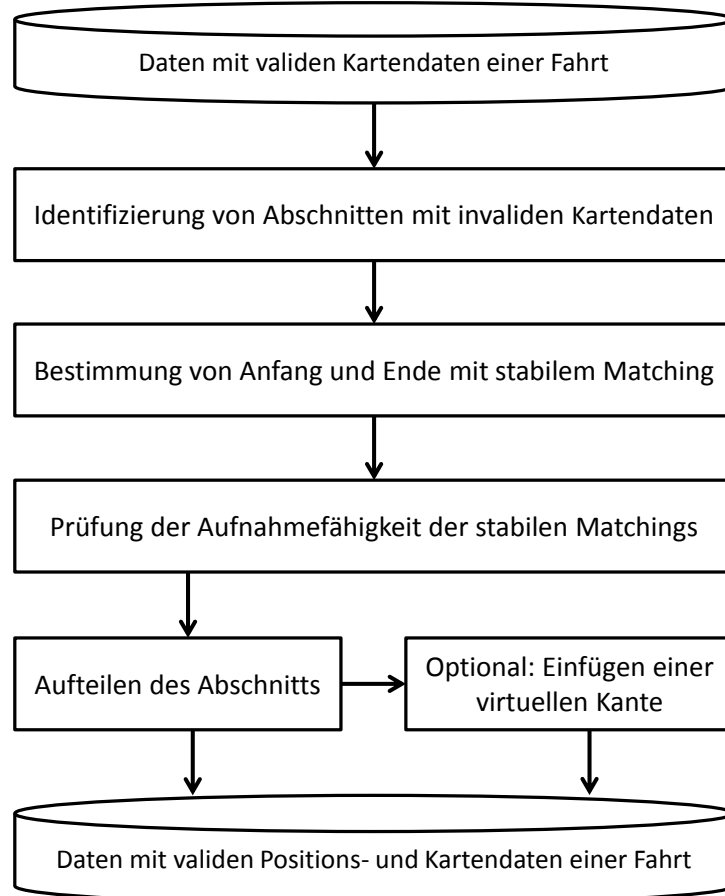


Abbildung 6.5: Vorgehen zur Bereinigung der Kartendaten

Übersteigt die gemessene die theoretische Länge, so werden die gematchten Kanten aufgefüllt und für die restliche Lücke eine *virtuelle Kante* mit einer eigenen ID und validen Kanteninformationen eingefügt. Auf diese Weise kann eine solche Fahrt weiterhin genutzt werden. Dieser Fall tritt insgesamt 390 Mal auf, so dass entsprechend viele virtuelle Kanten hinzugefügt werden.

In Abbildung 6.6, links ist ein Beispiel dargestellt, in dem eine Lücke zwischen zwei validen Kanten aufgeteilt wird. Auf der rechten Seite ist das Einfügen einer virtuellen Kante dargestellt. Auf diese Weise lassen sich die Kartendaten bereinigen und so ein durchgehend valide Kartendaten erzeugen.

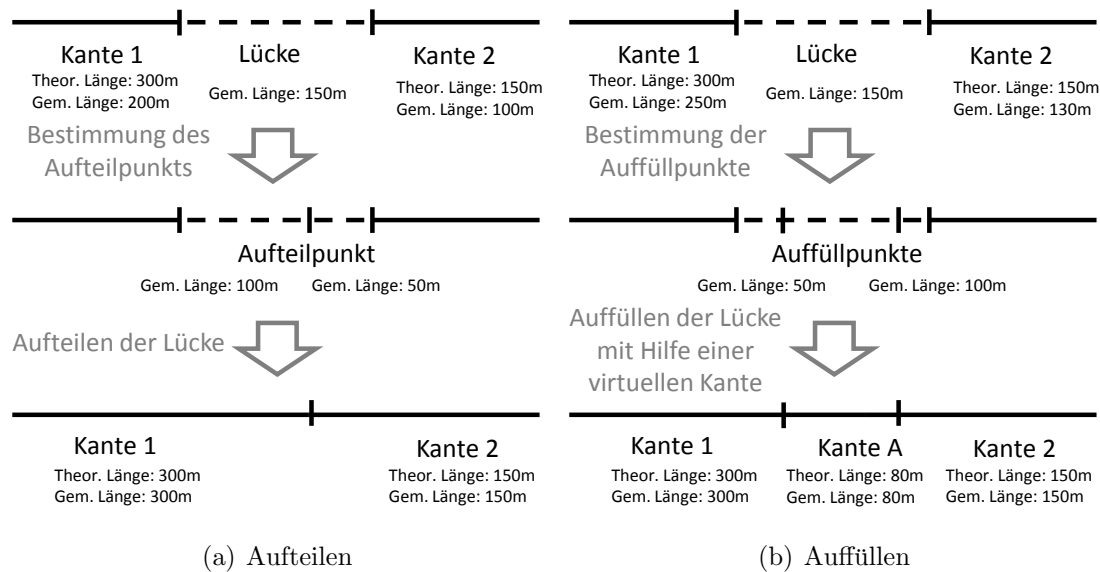


Abbildung 6.6: Beispiele für die Bereinigung der Kartendaten

6.2.3 Sprünge valider Positionsdaten

Im nächsten Schritt werden Sprünge in den Positionsdaten identifiziert. Besonders zu Beginn einer Fahrt kann es zu einem Sprung aufgrund der Justierung der Messtechnik kommen. Dabei kann zum einen die Position von invaliden auf valide Werte springen. Zum anderen kann es, z.B. aufgrund von Reflexionen, eine zu Beginn an sich valide Position korrigiert werden.

Zusätzlich werden Sprünge in den validen Positionsdaten gesucht. Das Kriterium ist dabei ein Sprung von mehr als hundert Metern. Bei Störungen kann es vorkommen, dass eine valide Position für längere Zeit gehalten wird und dann sprunghaft korrigiert wird, was zu diesen Ereignissen führt.

Das weitere Vorgehen ist in Abbildung 6.7 dargestellt. Bei einem Sprung von invaliden auf valide Werte am Anfang wird versucht, die Informationen nach dem Sprung nach vorne zurückzuschreiben. Dabei werden die in den vorherigen Schritten beschriebenen Techniken angewendet. Ein Sprung kommt bei sehr vielen Fahrten vor, doch findet dieser im Durchschnitt nach zwei Sekunden statt. Daher können die Daten meist überschrieben werden. Ist dieses nicht möglich, so wird die Fahrt bis zum ersten stabilen Matching abgeschnitten.

Bei einem Sprung mit validen Werten muss geprüft werden, ob es sich um einen doppelten Sprung handelt. In diesem Fall wird geprüft, ob ein Vorgehen wie bei invaliden Werten möglich ist. Bei einem solchen Doppelsprung sollten zwei Ketten

zusammenpassen und die dazwischenliegende Kette wird nach Möglichkeit linear analog zum Vorgehen bei Lücken interpoliert. Dies kann bei einer zwischenzeitlichen Fehlortung vorkommen.

Bei einem einfachen Sprung und der Bestimmung der vorherigen und nachfolgenden Sprüngen in der Position wird diese dazwischen interpoliert. Meist geht ein Sprung mit einem Wechsel der Kante einher. Zwar kann aufgrund der fehlenden Information nicht evaluiert werden, ob die Zuordnung zur Kante korrekt ist, doch kann durch eine Interpolation der Position zumindest sichergestellt werden, dass die Position vom Ende der vorherigen und Anfang der neuen Kante übereinstimmen.

Bei einigen Fahrten können kein stabiles Matching und damit keine valide Position bestimmt werden. Diese insgesamt 42 Fahrten werden von den weiteren Bereinigungsschritten ausgeschlossen. Das Ergebnis ist, dass in jeder Fahrt eine valide Position vorliegt und zu jedem Messpunkt Kartendaten verfügbar sind.

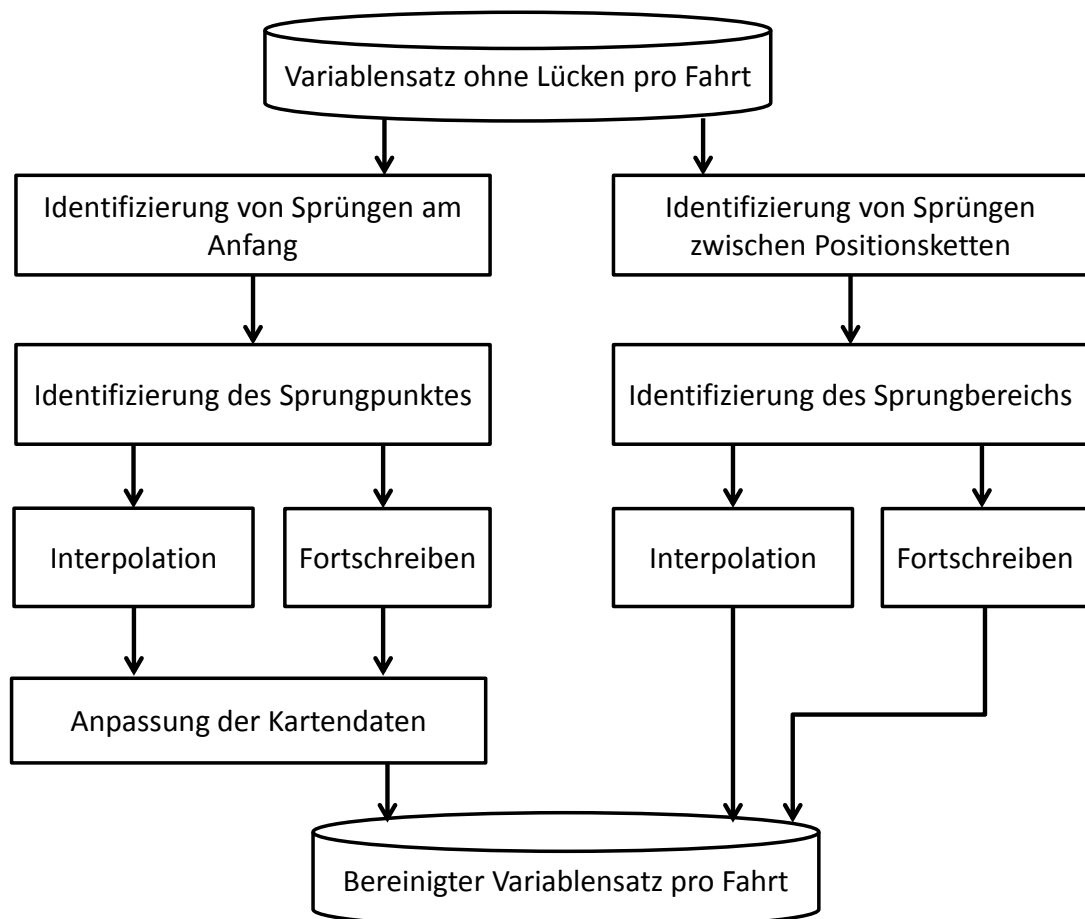


Abbildung 6.7: Vorgehen zur Bereinigung von Sprüngen

6.2.4 Matching

Da die Zuordnung der Position zu Kanten durch das Matching von entscheidender Bedeutung ist, wurde in dieser Arbeit das Matching erneut nachträglich verbessert und stellt den letzten Schritt der Datenbereinigung dar. In Abbildung 6.8 ist der Ablauf des Verfahrens dargestellt.

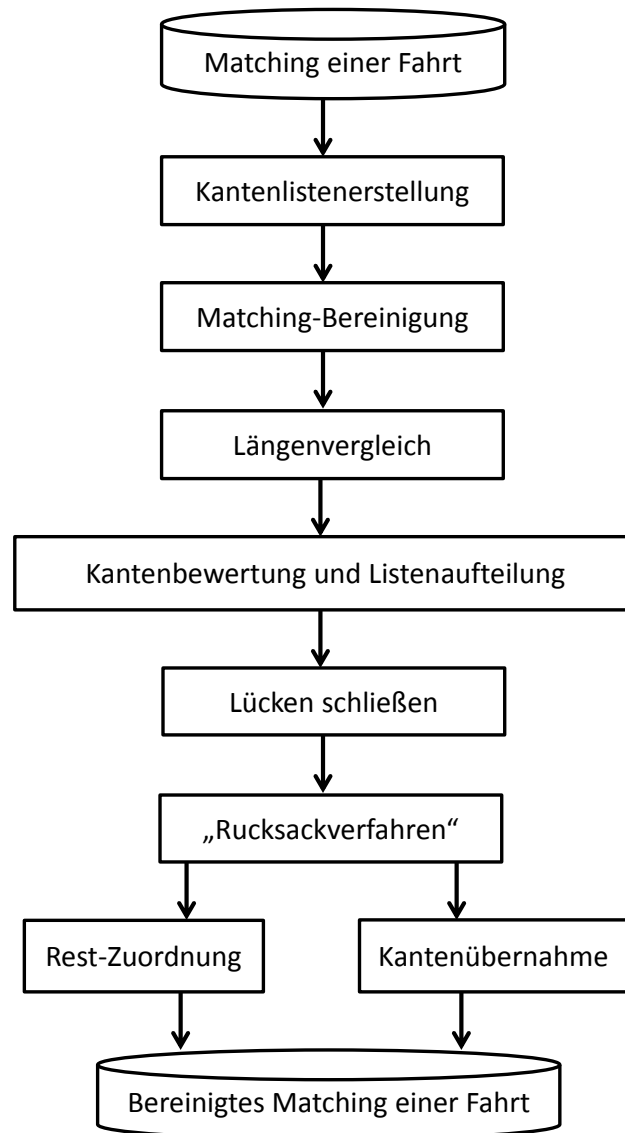


Abbildung 6.8: *Ablauf der Matchingverbesserung*

Zusätzlich zur Verbesserung der Ortung wird das Matching, also die Zuordnung einer Position zu einer Kante der digitalen Karte, bereinigt. Der im Projekt euroFOT verwendete rudimentäre Matching-Algorithmus, der die grundlegenden Informationen zur Analyse bereitstellte, wurde bereits im Projekt durch Plausibilitätsanalysen

verbessert (z.B. Autobahnbrücken). Die Zuordnung auf Basis dieser Korrekturen war für die Analysen im Projekt ausreichend, da hier hauptsächlich allgemeine Charakteristika, wie. z.B. die Straßenklasse, für die Analyse genutzt wurden.

Zuerst wird für eine Fahrt eine Liste der gematchten Kanten erstellt. Die Liste umfasst Anfangs- und Endzeitpunkt und somit die Dauer des Matchings auf der Kante sowie die zurückgelegte Entfernung in diesem Zeitraum. Dazu werden die Informationen über die Kante aus der digitalen Karte angehängt. Ziel des Verfahrens ist es, die Kanten mit gesicherten Matchings zu identifizieren und unplausible Matchings zu bereinigen.

Als erster Schritt werden analog zur Bereinigung der Daten aus dem vorherigen Abschnitt Doppelsprünge identifiziert, in denen ein Matching einer Position auf einer Kante kurzzeitig auf eine andere Kante springt und danach gesichert wieder zurück. Ein klassisches Beispiel für solch einen Sprung ist erneut eine Brücke. Diese Matching-Bereinigung verringert die Liste um kurze unplausible Matchings.

Für ein gesichertes Matching müssen die Kanten mindestens 50 Meter oder bei kurzen Kanten 90% ihrer theoretisch möglichen Länge Positionen zugeordnet sein. Andernfalls werden diese Kanten als falsch gematcht klassifiziert. Die anderen Kanten werden einer nächsten Prüfung bezüglich der Länge unterzogen. Dafür werden die, in der digitalen Karte gespeicherten, theoretischen Längen der Kanten herangezogen und mit den gemessenen Längen verglichen. Idealerweise sollten diese übereinstimmen. Kanten, deren gematchte Länge und die theoretisch vorgegebene maximal um 50 Meter oder 10% der Länge abweichen, werden als gesichert gematcht klassifiziert. Diese Kanten werden nicht verändert und teilen die Liste der Kanten einer Fahrt in Abschnitte auf. Für die anderen Kanten werden zunächst noch Sprünge in den Matchings mit längerer Matchinglänge gesucht. Dazu muss vor und nach einer unmarkierten Kante dieselbe Kante für mindestens 50 Meter gematcht worden sein. In diesem Fall wird die unmarkierte Kante der doppelt gematchten Kante zugeordnet und die drei Kanten verschmelzen zu einer Kante.

Im nächsten Schritt wird für jeden Abschnitt nun versucht, ein von der Länge möglichst gut passendes Matching zu erreichen. Das Verfahren verläuft analog zu bekannten *Rucksackproblemen* [46], wobei hier die Reihenfolge berücksichtigt wird. Kanten mit einem möglichst hohen Vergleichswert unterhalb der theoretisch vorgegeben Länge von über 50% werden markiert. Für diese markierten Nachbarn werden benachbarte unmarkierte Kanten gesucht, deren Matching auf die markierte Kante übertragen wird. So werden die Längen der Kanten nach Möglichkeit erhöht, die bereits einen hohen Übereinstimmungswert haben. Dieses Verfahren

wird so lange fortgesetzt bis keine unmarkierten Kanten mehr vorhanden sind oder die markierten Kanten auf einen Längenwert von maximal 110% der theoretischen Länge angereichert wurden. Bleiben nach dem Verfahren noch unmarkierte Kanten über, wird geprüft, ob diese eine weitere Kante bilden sollen, oder ob sie einer benachbarten Kante zugeschlagen werden. Dies hängt vom Absolutabstand der zwei Möglichkeiten von 100% der theoretischen Länge ab. Ein vereinfachtes Beispiel des Ablaufs des gesamten Verfahrens ist in Abbildung 6.9 dargestellt.

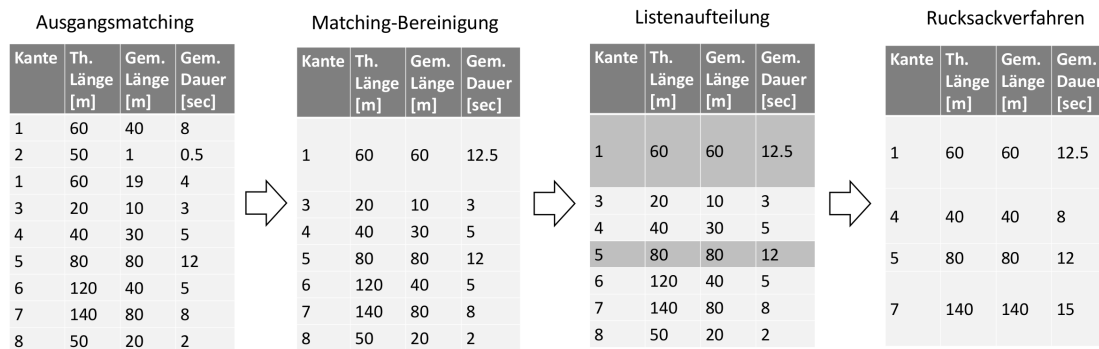


Abbildung 6.9: Visualisierung der Matchingverbesserung

Das Ergebnis des Verfahrens ist eine Kantenliste für jede Fahrt mit einem gesicherten und plausiblen Matching der Kanten, wie sie für die weitere Arbeit benötigt wird. Eine Visualisierung des Ergebnisses für einen Abschnitt von ca. 60 Sekunden ist in Abbildung 6.10 dargestellt.

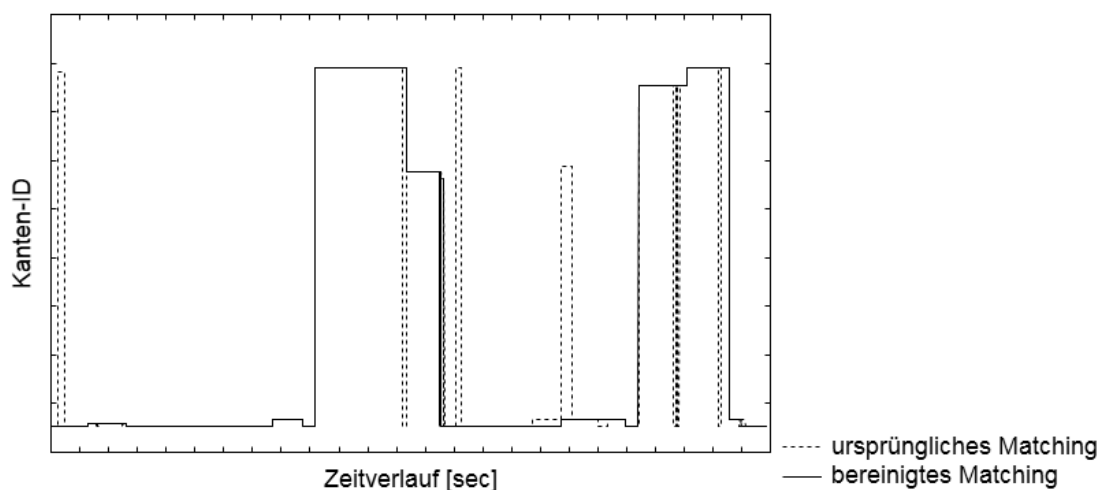


Abbildung 6.10: Visualisierung des Ergebnisses der Matchingverbesserung

Die gestrichelte Linie stellt die ID der ursprünglich gematchten Kanten dar, die durchgezogene entsprechend die ID der bereinigten Kanten. Deutlich sind kurze Sprünge aber auch längere Abschnitte zu erkennen, die bereinigt wurden.

6.3 Datenreduktion

Im nächsten Schritt werden die Daten auf die verschiedenen Datensätze, wie in Abbildung 6.1 dargestellt, reduziert. Das Ergebnis sind die Datensätze für die Analysen der Kanten-, Routen- und Netzwerkperspektive sowie die erstellte digitale Karte, die für die interne Routingreferenz verwendet wird.

6.3.1 Grundlage für die Kantenperspektive

Das Ergebnis der Datenbereinigung war eine Kantenliste für jede Fahrt eines Fahrers mit der Dauer sowie der auf der Kante zurückgelegten Entfernung. Aus dieser Kantenliste werden die Kanten, die eine Mindestlänge von fünf Sekunden haben, extrahiert. So werden nur robuste Messungen der Kanten genutzt. Diese Messungen werden über alle Fahrten pro Fahrer zusammengefasst. Das Ergebnis ist eine Menge an Messungen auf den extrahierten Kanten. Diese umfassen die Kantencharakteristika wie ID, theoretische Länge oder Geschwindigkeitsklasse sowie die gemessene Länge und Dauer.

Die Kanten-ID einer Messung wird für die weitere Reduktion nicht berücksichtigt. Somit können Kanten verschieden häufig in die Menge eingehen. Für die meisten Kanten gibt es allerdings, wie in Abbildung 6.11 für einen typischen Fahrer dargestellt, nur eine oder zwei Messungen. Die Anzahl der Messungen (*Benutzungsgrad*) nimmt danach deutlich ab. So gibt es nur noch ca. 300 Kanten mit einem Benutzungsgrad von mehr als 20.

Eine ID-basierte Mengenerstellung wäre zwar möglich, benötigt allerdings einen deutlich höheren Rechen- sowie Speicherbedarf. Zudem verläuft das hier dargestellte Vorgehen analog zu existierenden Verfahren und ermöglicht somit eine Vergleichbarkeit der Analyseergebnisse für die Kantenperspektive.

In Abbildung 6.12 sind die Anzahl an Messungen bzw. im Folgenden ID-unabhängigen Kanten pro Fahrer dargestellt. Der Durchschnitt liegt bei über 20 000 Kanten, wobei es große Unterschiede zwischen den Fahrern gibt. So liegt das Maximum bei über 70 000 Kanten. Das Minimum liegt bei 10 000 Kanten, was eine genügend große Menge an Kanten für die folgenden Analysen darstellt. Die Menge der Kan-

ten wird abschließend in Untermengen in Abhängigkeit der Geschwindigkeitsklasse unterteilt. Diese bilden die Grundlage für die Analysen der Kantenperspektiven.

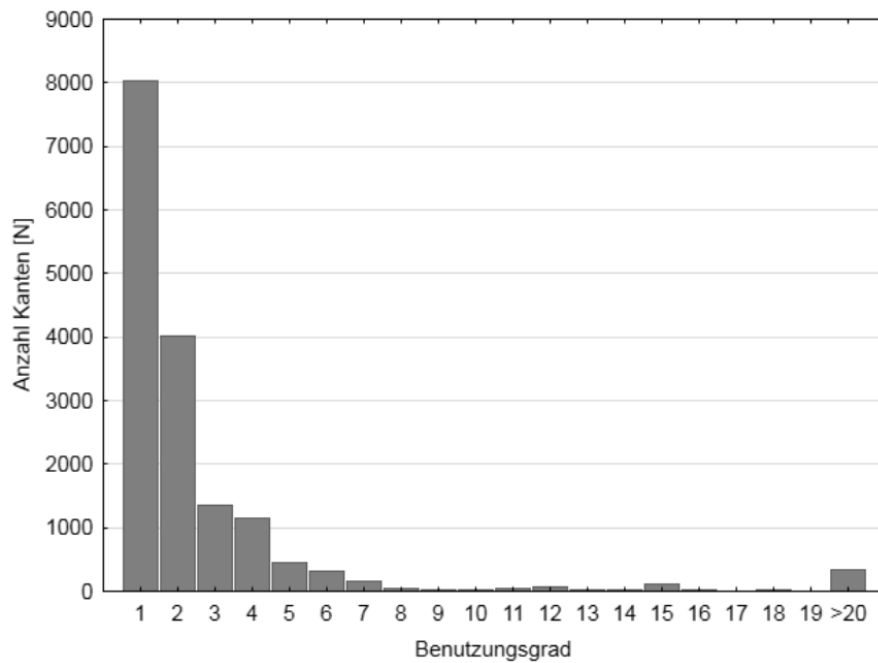


Abbildung 6.11: *Häufigkeitsverteilung des Benutzungsgrads der Kanten eines Fahrers*

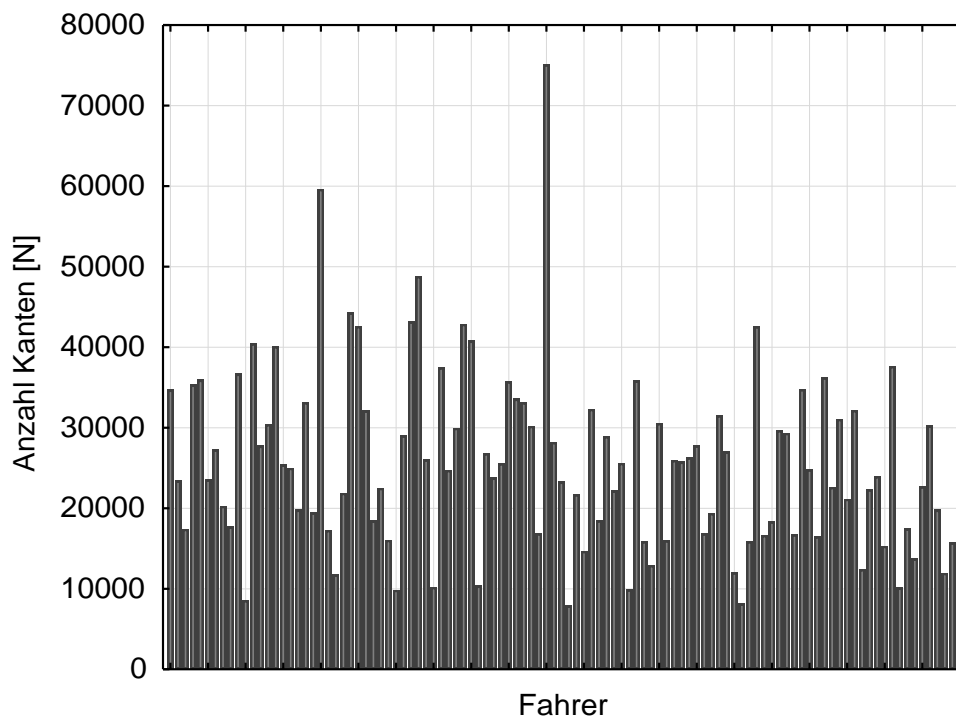


Abbildung 6.12: *Anzahl der ausgewählten Kanten eines Fahrers*

6.3.2 Grundlage für die Routenperspektive

Für jeden Fahrer werden eine Liste der FR-Routen und deren Charakteristika erstellt. Diese Liste wird auf die Routen mit einer Länge von mindestens einem Kilometer reduziert. Auf diese Weise werden nur Routen in die Analysen der Routenperspektive einbezogen, die eine größere Anzahl an Kanten des Straßennetzes nutzen.

Für die reduzierten Routen werden Start- und Zielpunkte zu Beginn und Ende der Fahrt bestimmt. Dazu wird der Routenverlauf für jede Route gespeichert. Eine Visualisierung einer Route ist in Abbildung 6.13 zu finden. Diese illustriert Start- und Zielpunkt sowie Verlauf der Route. Zusätzlich werden Länge und Dauer der Route einer Fahrt berechnet und in der Liste abgelegt.

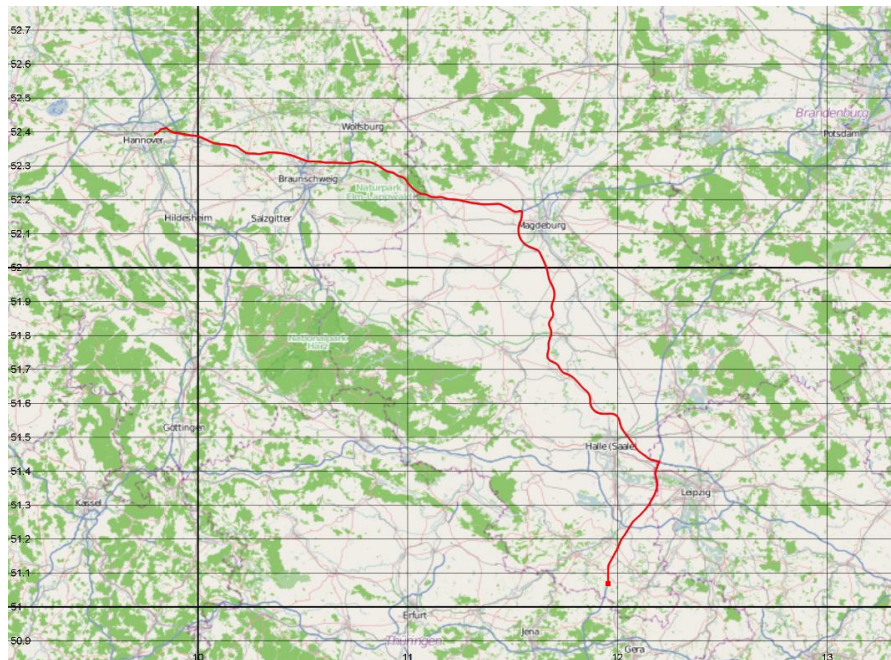


Abbildung 6.13: *Visualisierung einer Route*

Das Ergebnis ist ein Datensatz mit einer Routenliste für jeden Fahrer, sowie den Charakteristika und den Positionen der Routenverläufe für jede Route. In Abbildung 6.14 sind die Anzahl der Routen pro Fahrer dargestellt. Ähnlich wie im vorherigen Abschnitt gibt es große Unterschiede zwischen den Fahrern. Im Mittel können 230 Fahrten pro Fahrer für die Analyse der Routen genutzt werden. Das Minimum liegt bei 70 Routen.

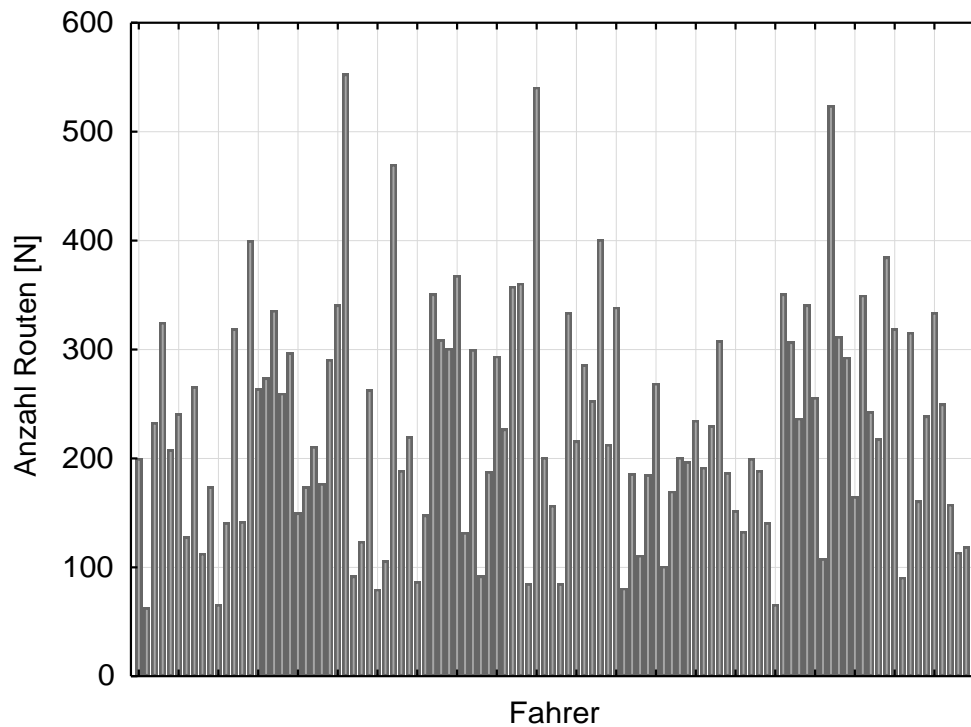


Abbildung 6.14: Anzahl FR-Routen länger als 1 km pro Fahrer

6.3.3 Grundlage für die Netzwerkperspektive

Für die Netzwerkperspektive wird analog zum Vorgehen für die Kantenperspektive eine Kantenliste über alle Fahrten eines Fahrers erstellt. Diese ist Kanten-ID-orientiert. Jede Messung wird einer Kante, die eindeutig durch die ID bestimmt wird, zugeordnet. Damit kann jede Kante mit eindeutiger ID mehrere Messungen enthalten.

Von den theoretischen in der digitalen Karte gespeicherten Werten wird außer der ID noch die Richtung der Kante genutzt. Sie bestimmt bei Benutzung der Kante welche Position der Start- und welcher der Zielposition zugeordnet wird. Für die anschließende Erstellung des Graphen wird die Richtung nicht weiter berücksichtigt. Für jede Messung auf einer Kante wird der Benutzungsgrad auf der Kante erhöht und bestimmt auch die Anzahl an gemessenen Start- und Zielpositionen. An dieser Stelle sei anzumerken, dass die GPS-Position des Anfangs und Endes der Kante in euroFOT nicht aus der digitalen Karte extrahiert wurde. Dazu wird noch die adjazente Kante gespeichert, d.h. die Kante, die zu der aktuellen Kante benachbart ist. Auch die Anzahl der adjazenten Kanten entspricht der Anzahl der Messungen, wobei es zu Mehrfachmessungen einer adjazenten Kante kommen kann.

Das Ergebnis ist eine Kantenliste pro Fahrer, die einen Benutzungsgrad sowie entsprechend viele Messungen der Start- und Zielposition sowie adjazente Kanten beinhaltet. Für die weitere Reduktion wird der Schwerpunkt auf den Zusammenhang sowie die Anfangs- und Endposition der Kanten gelegt. Die Positionen bilden Knoten. Der Zusammenhang wird zum einen durch den logischen Zusammenhang, abgeleitet aus dem Routenverlauf, gebildet. Zum anderen auch auf Basis der Position. Knoten die nahe beieinanderliegen, bspw. komplexe Kreuzungen, werden zu einem Superknoten zusammengefasst. Das Verfahren ist in Abbildung 6.15 dargestellt.

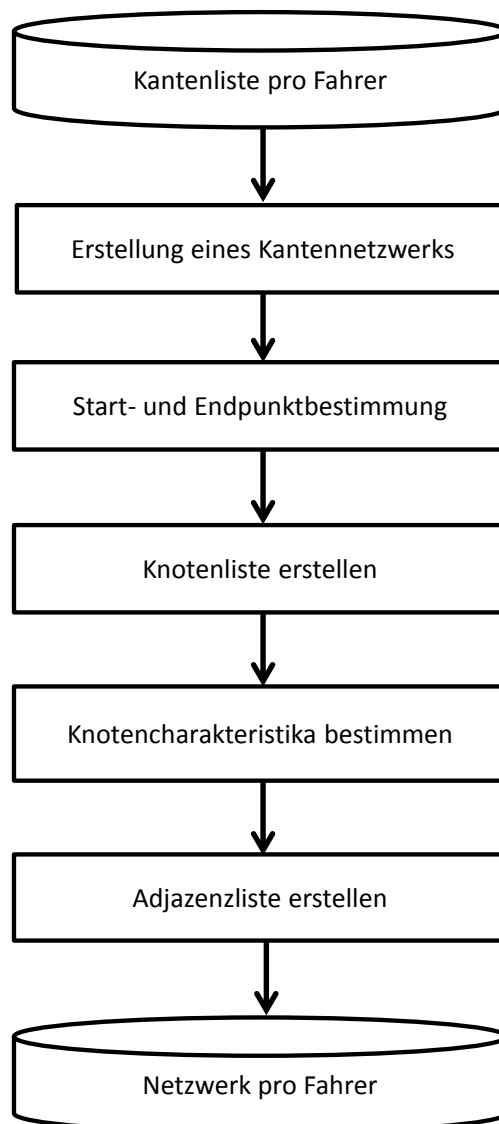


Abbildung 6.15: *Ablauf der Netzwerkerstellung*

Zunächst müssen die Positionen der Start- und Endpunkte der Kanten bestimmt werden, wofür die Richtung der Kante einbezogen wird. Ist eine Kante nur einmal genutzt worden, werden die Positionen übernommen. Bei Mehrfachbenutzung wird der Median berechnet. Der Median hat gegenüber dem Mittelwert den Vorteil, dass ein Ausreißer besser aufgefangen werden kann. Dazu ist die genutzte Position eine in der Realität gemessene und sollte somit auf und nicht zwischen dem Straßennetz liegen. Die so bestimmten Positionen werden nun für die Erstellung einer positionsbasierten Knotenliste genutzt. Dazu wird in der bereits existierenden Liste jeweils für den Anfangs- und Endpunkt einer Kante nach einem Knoten gesucht, der weniger als 50 Meter entfernt liegt. Wird so ein Knoten gefunden, so wird der Knoten mit der entsprechenden inzidenten Kante und dem adjazenten Knoten in die Knotenliste an der Stelle des bereits vorhandenen Knotens aufgenommen.

Wird kein Knoten gefunden, so bildet die Position einen neuen Knoten. Bei einer nahtlosen Datenaufzeichnung sollten die Positionen der Endpunkte einer Kante und der Anfangspunkt der neuen Kante gleich sein und diese Punkte damit zum gleichen Knoten gehören. Für die letzte Position einer Route und die erste der nächsten Fahrt sollte dies ebenfalls häufig der Fall sein. Durch die positionsgestützte Knotenbildung werden komplexe Kreuzungen zusammengefasst, allerdings auch nahe beieinander liegende Kreuzungen.

Daher kann es dazu führen, dass Knoten viele adjazente Knoten haben. Ebenfalls kann es vorkommen, dass Straßen, abgelegt in der digitalen Karte, sich nicht kreuzen (z.B. Brücken) nun durch einen Knoten verbunden sind. Aufgrund der Zielsetzung der Netzwerkanalyse, wie sie in Abschnitt 5.2.3 beschrieben wurde, ist dies allerdings vernachlässigbar. Vielmehr ist der strukturelle Zusammenhang entscheidend.

Mit diesem Verfahren lässt sich für jeden Fahrer eine Knotenliste erstellen, die ein Netzwerk beschreibt. Jeder Knoten des Netzwerks besitzt Charakteristika, wie die Position und eine Liste der inzidenten Kanten und der adjazenten Knoten. In Abbildung 6.16 ist eine grafische Veranschaulichung des Vorgangs zur Bildung des Netzwerks, das aus der Knotenliste und den entsprechenden Kanten besteht, dargestellt. Die Häufigkeit der Benutzung der Kanten ist farblich markiert. Dieses Netzwerk bildet die Grundlage für die Analysen für jeden Fahrer.

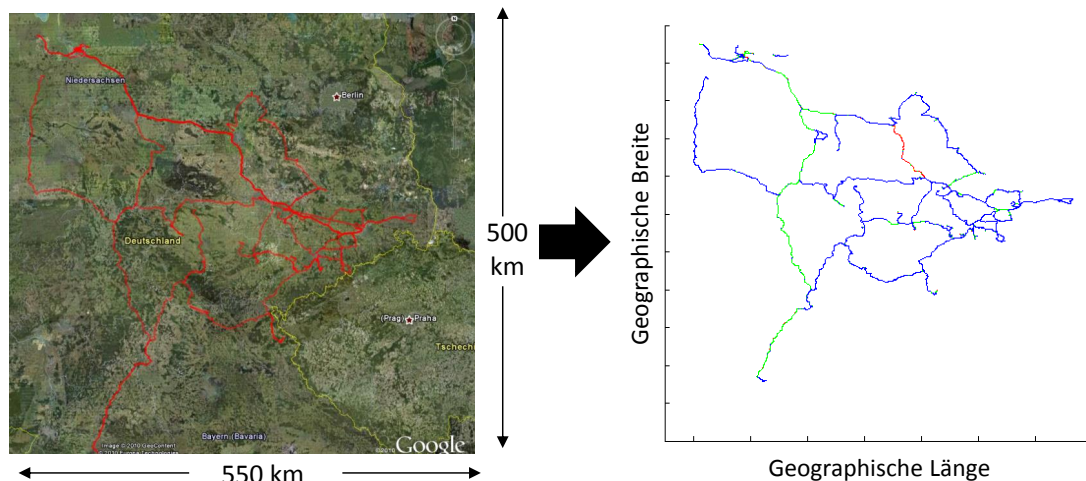


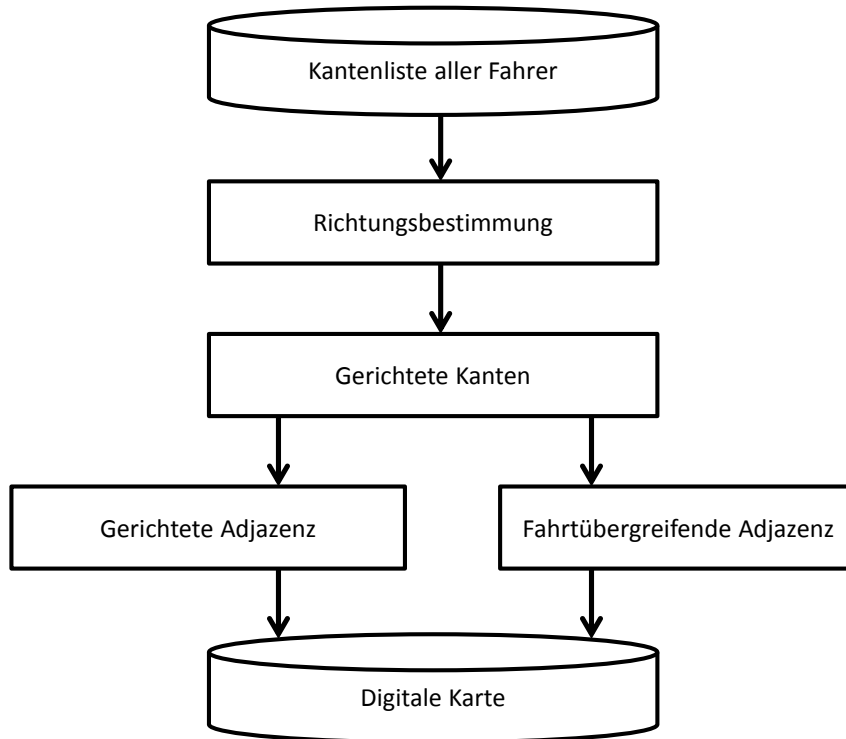
Abbildung 6.16: Visualisierung der Netzwerkerstellung, (Benutzungsgrad 1=blau, Benutzungsgrad 2=grün, Benutzungsgrad größer 2=rot)

6.3.4 Digitale Karte für die interne Routingreferenz

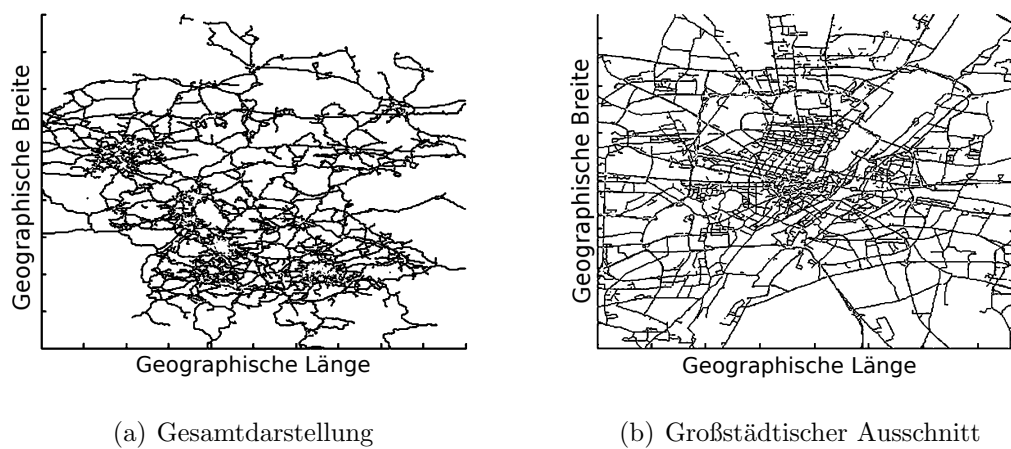
Zum Abschluss der Datenaufbereitung wird das Verfahren vorgestellt, das eine digitale Karte auf Basis der Kanten aller Fahrer erstellt, die für das spätere Routing der internen Referenz genutzt wird. Das Verfahren verläuft analog zur Netzwerkerstellung. Allerdings werden die Fahrten aller Fahrer zusammen betrachtet und der Schwerpunkt bei der Netzbildung auf den logischen Zusammenhang, wie er sich aus dem Verlauf der gemessenen Routen ergibt, gelegt. Damit ist der entsprechende Graph kanten- und nicht knotenbasiert.

Dadurch soll verhindert werden, dass Brücken als Kreuzungen in die Karte aufgenommen und bei der Routenberechnung genutzt werden, was die Ergebnisse verzerren könnte. Das Verfahren ist in Abbildung 6.17 dargestellt.

Die Kantenliste wird, wie beim vorherigen Verfahren, durchsucht. Da das so gebildete Netzwerk als Routinggrundlage dienen soll, ist die Richtung wichtig, um die Information zu erhalten, welche Kante auf die vorherige in welcher Richtung folgt. So werden z.B. Einbahnstraßen richtig behandelt. Eine in beide Richtungen befahrbare Kante wird so in zwei aufgeteilt (positiv und negativ). Auf Basis der eingetragenen Richtung wird die befahrene Kante zugeordnet. Die Adjazenz ergibt sich aus der folgenden Kante, die in der entsprechenden Liste aufgenommen wird, falls sie noch nicht vorhanden ist.

Abbildung 6.17: *Ablauf der Kartenerstellung*

Das Ergebnis ist eine gerichtete Adjazenzliste der gemessenen Kanten, die angibt welche Kante von welcher Kante erreichbar ist und bildet die Grundlage für die genutzte interne Routingreferenz. In Abbildung 6.18, links ist die auf diese Weise erstellte Karte nahezu komplett und rechts ein städtischer Ausschnitt zur Veranschaulichung dargestellt und zeigt, dass die Karte ein großes und verzweigtes Netz für die von den Fahrern genutzten Gebiete aufspannt.

Abbildung 6.18: *Visualisierung der erstellten Karte*

7 Empirische Untersuchungen zur modellgestützten Personalisierung

In diesem Kapitel werden die Ergebnisse der empirischen Untersuchungen der in Kapitel 6 aufbereiteten euroFOT-Daten für die Modellierung und die Personalisierung dargestellt und erläutert. Die Untersuchungen umfassen die Kanten- und die Routenanalyse sowie die Zuordnung der Fahrer zu einem Navigationstyp (Netzwerkperspektive). Daran schließt sich die Korrelationsanalyse sowie die Erstellung eines zweidimensionalen Modells der navigatorischen Präferenzen an. Die Analyse bildet die Grundlage für die Personalisierung, die durchgeführt und bewertet wird. Damit folgt der Aufbau dem in Kapitel 5 beschriebenen Verfahren. Abschließend werden die Ergebnisse diskutiert.

7.1 Ergebnisse für die Kantenperspektive

Die Analyse der Kantenperspektive umfasst zwei Schritte. Für jeden Fahrer wird pro Klasse eine angepasste Geschwindigkeit berechnet, um sie für die Personalisierung zu nutzen. Das Vorgehen verläuft analog zu bestehenden Verfahren in heutzutage verfügbaren Navigationssystemen, siehe Abschnitt 2.3.5. Dazu werden die Ergebnisse für das Modell 1 verwendet, um ein Maß M_{v_f} für jeden Fahrer f zu berechnen. Eine detaillierte Beschreibung des Vorgehens ist in Abschnitt 5.2.1 zu finden.

7.1.1 Anpassung der Kantengewichte für die Personalisierung

Zunächst wird die in der Datenaufbereitung erstellte Datenbasis für die Kantenanalyse bewertet. Dazu wird für jeden Fahrer die Anzahl an ID-unabhängigen Kanten der Geschwindigkeitsklassen analysiert. In Abbildung 7.1 sind die Ergebnisse als *Whisker-Box-Plots* mit Ausreißern dargestellt [12]. Der Punkt in den Boxen definiert den Median. Die Boxen werden durch das 25% und 75% Perzentil bestimmt.

Die Intervalle (Whisker) wiederum geben den Bereich ohne Ausreißer an. Dieser beinhaltet Werte, die maximal das Anderthalbfache des *Interquartilsabstands*, also der Länge der Box, von den Rändern der Box entfernt sind.

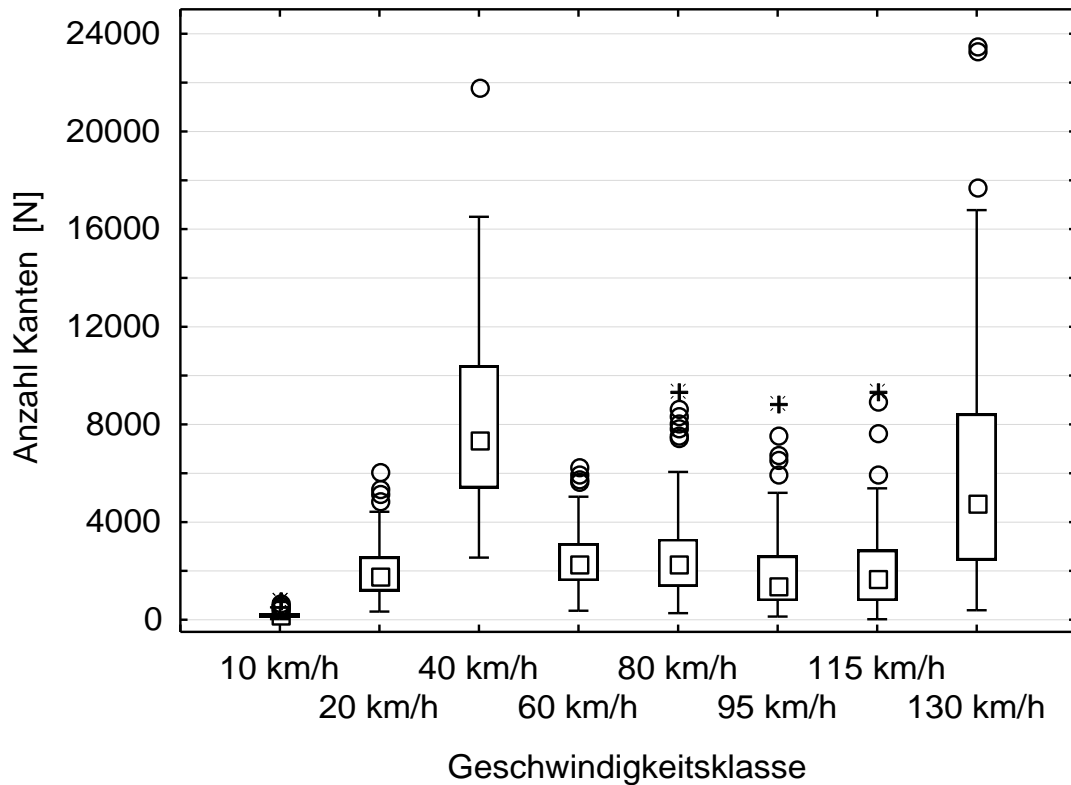


Abbildung 7.1: *Anzahl Kanten pro Geschwindigkeitsklasse für jeden Fahrer*

Es werden insgesamt acht Geschwindigkeitsklassen unterschieden. Die unteren drei Klassen (10 km/h, 20 km/h, 40 km/h) sind innerstädtische Haupt- und Nebenstrecken. Die mittleren Klassen (60 km/h, 80 km/h, 95 km/h) sind zumeist Überlandstrecken, während die beiden höchsten Klassen (115 km/h, 130 km/h) der Autobahn zugeordnet werden können. Die Unterschiede zwischen den Fahrern, die bereits in Abbildung 6.12 beobachtet wurden, zeigen sich auch bei der klassenbasierten Analyse. Dazu gibt es Unterschiede zwischen den Klassen. Für die meisten Fahrer bewegt sich die Anzahl der Kanten zwischen 1000 und 4000 pro Geschwindigkeitsklasse. Eine Ausnahme bildet die niedrigste Geschwindigkeitsklasse mit meist nur wenigen hundert Kanten.

Dazu sind die Klassen 40 km/h und 130 km/h hervorzuheben. Die Klasse 40 km/h beinhaltet Hauptverkehrs- und größere Nebenstraßen im Stadtgebiet. In dieser Klasse beträgt der Median über alle Fahrer 7500 Kanten bis hin zu einem Ma-

ximum von 22 000 Kanten für einen Fahrer. Die Anzahl in der höchsten Geschwindigkeitsklasse ist ebenfalls deutlich höher (Median 5 200 Kanten, Maximum 24 000 Kanten). Dies sind unbeschränkte Autobahnabschnitte. Damit sind insgesamt trotz der Einschränkung auf valide Kanten für alle Fahrer und für nahezu alle Klassen eine große Anzahl an Kanten für die Analyse vorhanden.

Im nächsten Schritt werden für jeden Fahrer pro Klasse neue Geschwindigkeiten für die Personalisierung berechnet. Dazu wird der Mittelwert der Geschwindigkeiten ermittelt. Das Ergebnis ist in Abbildung 7.2 dargestellt. Mit steigender Geschwindigkeitsklasse steigen auch die gemessenen Geschwindigkeiten an. Dabei wird auch die Streuung zwischen den Fahrern zunehmend größer. Eine Ausnahme bildet die Klasse 10 km/h mit einer ebenfalls großen Streuung.

Die Mediane über alle Fahrer entsprechen ab Klasse 40 km/h den angenommenen Geschwindigkeiten der Klasse. Dies zeigt, dass der generell angenommene Wert für die Geschwindigkeit in der digitalen Karte sinnvoll ist, veranschaulicht allerdings auch, dass aufgrund der Streuung zwischen den Fahrern eine Personalisierung von Nutzen sein und zu unterschiedlichen Routen für die Fahrer führen kann.

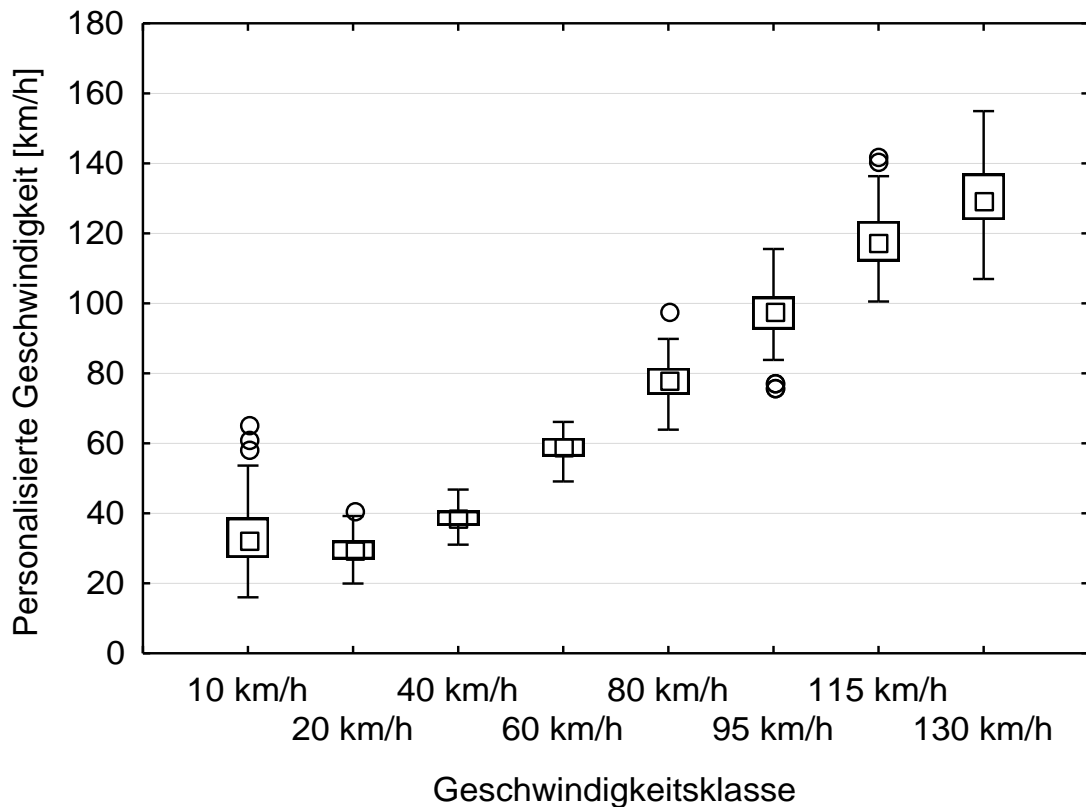


Abbildung 7.2: *Gemessene Geschwindigkeit pro Geschwindigkeitsklasse für jeden Fahrer*

7.1.2 Erstellung des Modells

Zur Bestimmung des Maßes M_v für die Kantenperspektive werden die für jeden Fahrer ermittelten Geschwindigkeiten mit denen in der digitalen Karte gespeicherten in Relation gesetzt. Zunächst wird in Abbildung 7.3 pro Fahrer aufgetragen, in welchen Klassen er schneller und in welchen er langsamer als die angenommene Geschwindigkeit ist. Zwar gibt es Fahrer, die generell immer schneller sind, doch liegt der Median pro Fahrer bei gut 60%, was fünf Klassen entspricht. Das Ergebnis zeigt deutlich, dass die meisten Fahrer generell weder schneller oder langsamer sind. Zudem sind verschiedene Fahrer auf verschiedenen Klassen langsamer oder schneller als die angenommene Geschwindigkeit.

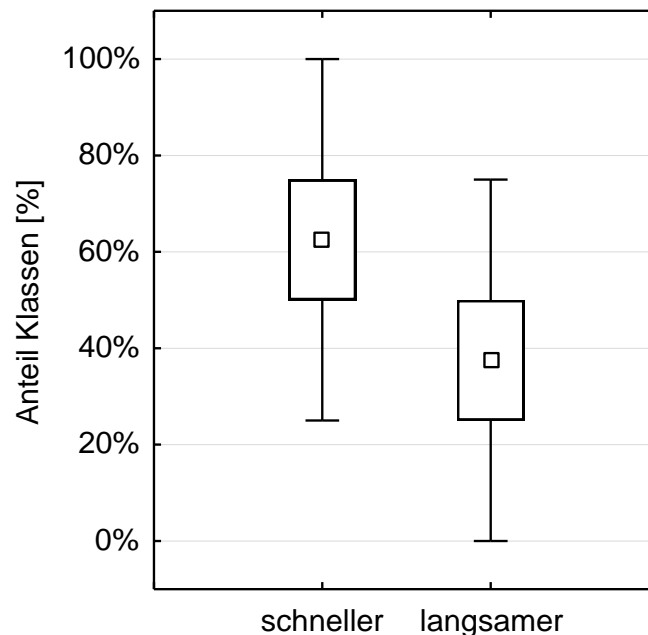


Abbildung 7.3: *Bewertung der gemessenen Geschwindigkeiten pro Fahrer*

Für eine Anpassung der Kantengewichte bedeutet dies, dass eine generelle oder straßenklassenbasierte Veränderung der angenommenen Geschwindigkeit, wie es in heutigen Navigationssystemen möglich ist, für diese Fahrer nicht sinnvoll ist. Vielmehr ist eine Differenzierung und direkte Anpassung der einzelnen Geschwindigkeitsklassen, wie sie im vorherigen Abschnitt durchgeführt wurde, sinnvoll. Für die Modellierung bedeutet dies ebenfalls, dass nicht nur eine Klasse herangezogen werden kann, die das Verhalten der Fahrer abbildet.

Für die Modellierung werden die Formeln aus Gleichung 5.3 in Abschnitt 5.2.1 angewendet. Das Ergebnis in Abbildung 7.4 zeigt den 0-Relativ-Unterschied für alle

Fahrer in den jeweiligen Klassen an. Für die Klassen ab 40 km/h weisen die Werte für alle Fahrer eine Abweichung zwischen -20% und 20% auf. Für die Geschwindigkeitsklassen 10 km/h und 20 km/h sind allerdings alle Werte positiv.

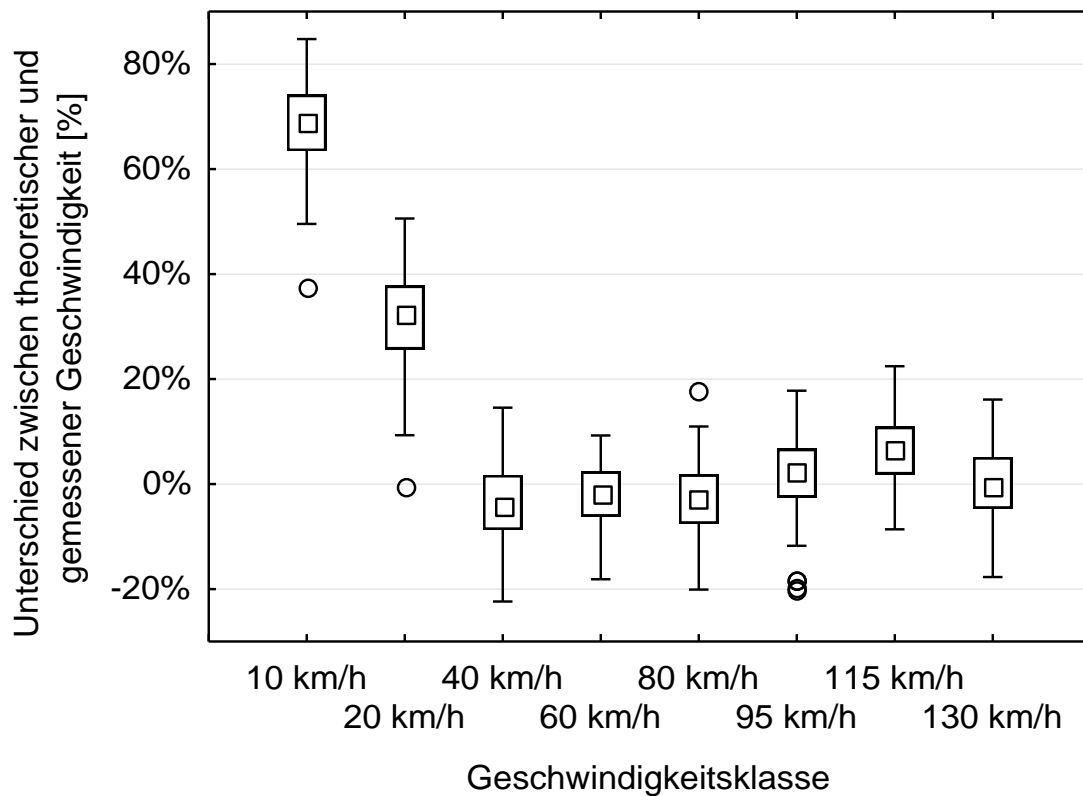


Abbildung 7.4: *Unterschied (0-Relativ-Vergleich) zwischen theoretischer und gemessener Geschwindigkeit pro Geschwindigkeitsklasse für jeden Fahrer*

In den Klassen mit mittlerer Geschwindigkeit ist das Verhältnis relativ ausgeglichen mit einer leichten Tendenz dazu, dass die Fahrer eher langsamer als angenommen sind. Für die Klassen 90 km/h und 115 km/h ist der Anteil schnellerer Fahrer erhöht. In der höchsten Klasse ist das Verhältnis wieder nahezu ausgeglichen.

Aufgrund der geringen Mächtigkeit und der starken Abweichung wird die Geschwindigkeitsklasse 10 km/h für die Modellierung nicht berücksichtigt, um eine Verzerrung des Ergebnisses zu vermeiden. Alle anderen Klassen mit validen Werten gehen mit gleichem Gewicht ein. Somit wird für die Kennwertbestimmung M_{v_f} für jeden Fahrer f der Modellierung der Mittelwert über alle validen Klassen ermittelt. Das Ergebnis ist in Abbildung 7.5 pro Fahrer aufgetragen.

Generell haben die Fahrer einen positiven Kennwert für M_v . Dies liegt zum einen an dem relativ hohen Wert für die Geschwindigkeitsklasse 20 km/h aber auch an

dem Ergebnis, dass es mehr Klassen gibt, in denen der Fahrer schneller ist, siehe Abbildung 7.3. Hier wird dieser Sachverhalt exakter dargestellt. Dazu kommt es zum Teil zu großen Unterschieden zwischen den Fahrern. Die Kennwerte M_{v_f} pro Fahrer für das Modell 1 sind der Dimension der Zeit zugeordnet und gehen in das zweidimensionale Modell ein.

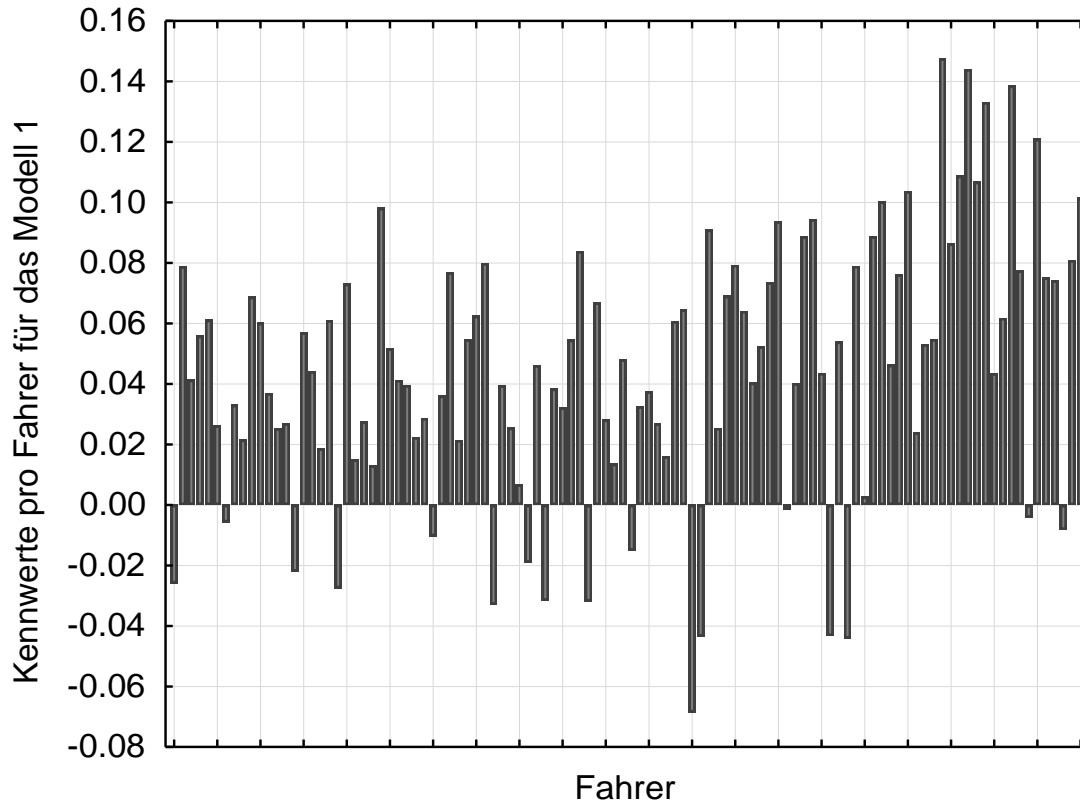


Abbildung 7.5: *Ergebnis der Modellierung der Kantenperspektive pro Fahrer*

7.2 Ergebnisse für die Routenperspektive

Die Modellierung für die Routenperspektive mit den Maßen M_d und M_l erfolgt mit Hilfe von Benchmarkrouten, die mit den Fahrerrealrouten verglichen werden, siehe Abschnitt 5.2.2 für eine detaillierte Beschreibung. Nach den Datenaufbereitungsschritten verbleiben ca. 24400 Routen, die die Kriterien erfüllen, dass der Fahrer die Route selbstständig im bekannten Gebiet gewählt hat. Für die Start-Zielkombinationen dieser Routen wird mit Hilfe der externen Routingreferenz eine schnellste Route berechnet. In Abbildung 7.6 ist die Anzahl der berechneten Routen für jeden Fahrer dargestellt.

Wie bereits in Abbildung 6.14 gesehen, gibt es eine große Streuung zwischen den Fahrern bei der Anzahl der gemessenen Routen. Entsprechend ist auch bei den berechneten Routen eine große Streuung vorhanden. Der Mittelwert liegt bei 220 berechneten Routen, wobei diese Zahl auch deutlich größer sein kann (Maximum 570). Das Minimum liegt bei 60 Fahrten, die für die Modellierung genutzt werden.

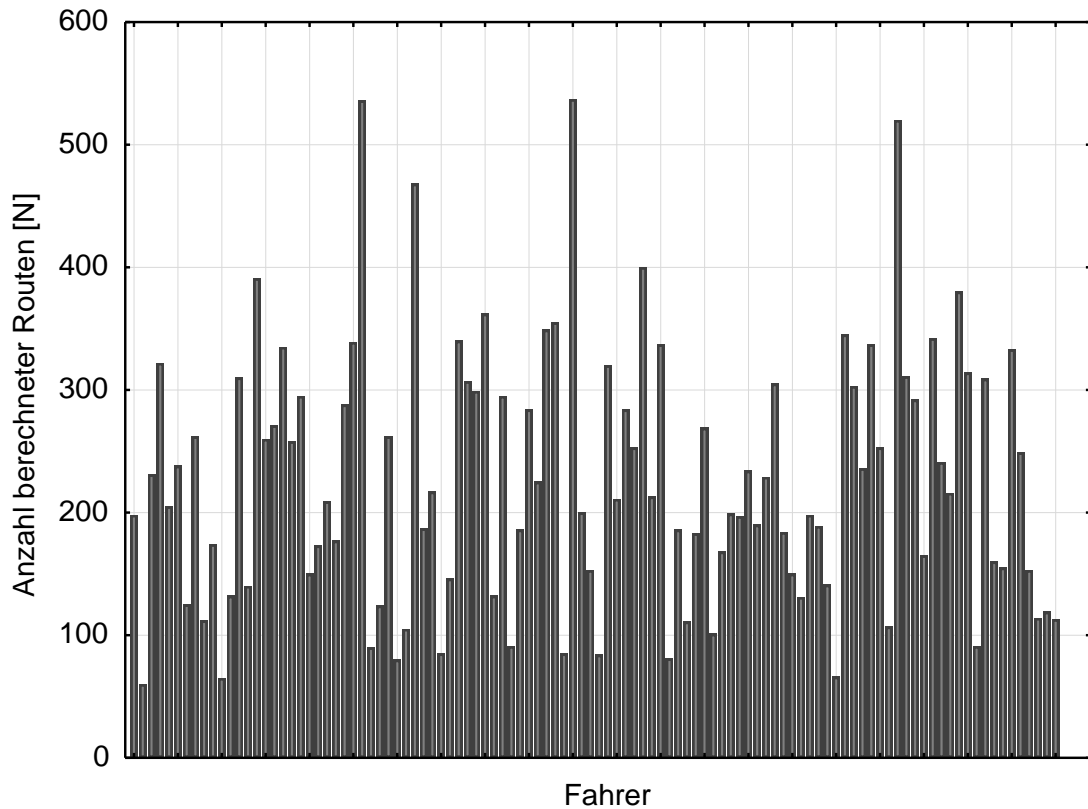


Abbildung 7.6: *Anzahl berechneter Routen*

Diese Anzahl lässt sich nur marginal erhöhen, wie Abbildung 7.7 zeigt. In der Abbildung ist pro Fahrer der relative Anteil aufgetragen, für die die Berechnung erfolgreich war. Für fast alle Fahrer konnten mehr als 95% der Routen berechnet werden. Nur bei zwei Fahrern sinkt der Wert auf 65%, wobei diese Fahrer eine hohe absolute Anzahl aufweisen. Gründe für eine nicht erfolgreiche Berechnung können ein ungenügend korrigiertes Matching aber vor allem auch inhaltliche Gründe, wie eine nicht zugangsberechtigte Straße bzw. eine Straße, die nicht in der digitalen Karte hinterlegt ist, sein.

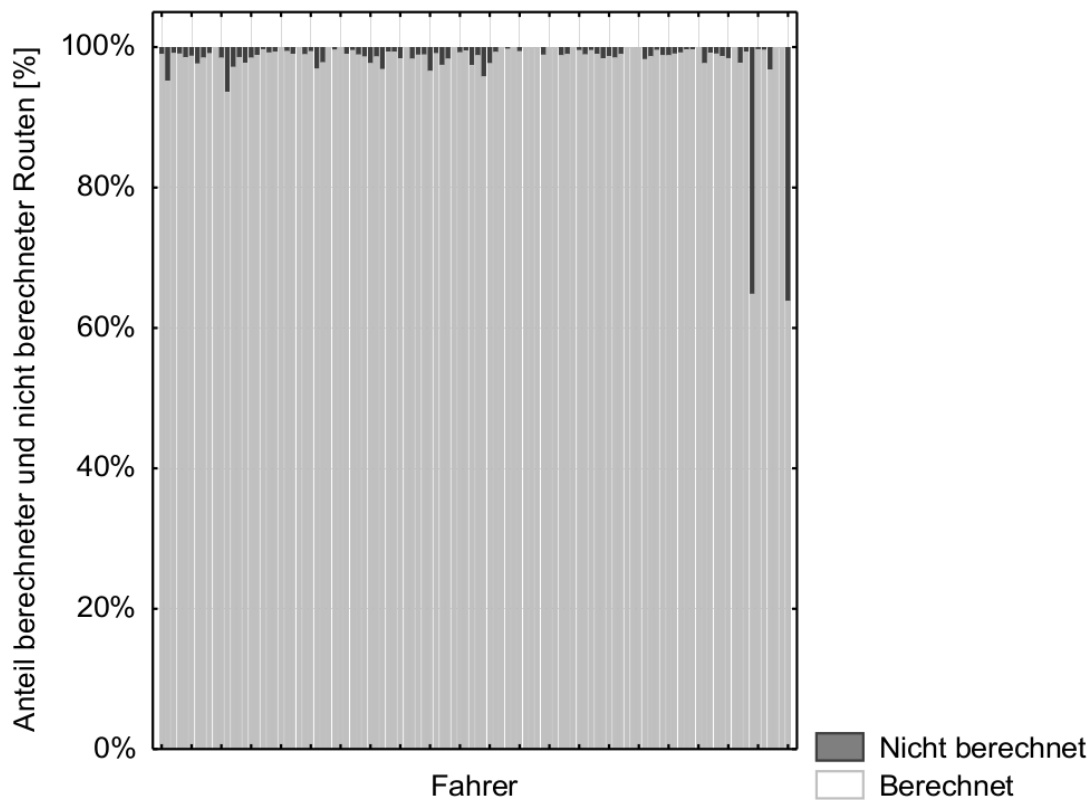


Abbildung 7.7: Anteil berechneter Routen

Für die Bestimmung von M_d und M_l wird zunächst die relative Abweichung der Routendauer und Routenlänge der Fahrerrealrouten von den Referenzrouten berechnet, siehe Gleichungen 5.4 und 5.5. Daran anschließend werden die Abweichungen für die Fahrer zusammengefasst und der Median als robustes Maß gebildet. Damit geht jede Route mit gleichem Gewicht in die Berechnung ein.

Das Ergebnis ist in Abbildung 7.8 dargestellt. Auf der X-Achse sind die Kennwerte M_{l_f} und auf der y-Achse die Kennwerte M_{d_f} aufgetragen. Jeder Punkt repräsentiert einen der Fahrer. Ein positiver Wert bedeutet, dass ein Fahrer über alle seine Routen schnellere bzw. kürzere Routen als die von der Referenz berechneten Route ausgewählt hat. Ist dieses beim Kriterium der Länge durchaus plausibel, so ist es beim Kriterium der Dauer überraschend, dass Fahrer einen positiven Wert erreichen und damit über alle Fahrten schneller als die schnellste berechnete Route sind. Auch ist die Streuung beim Vergleich der Maße der Dauer deutlich höher als beim Vergleich der Maße der Länge. Dies liegt daran, dass die Dimension der Zeit mehr externen Einflüssen, wie z.B. Stau oder aber auch die Wahl der Geschwindigkeit, ausgesetzt ist. Dementsprechend ist auch die Streuung zwischen den Fahrern größer.

Ebenso sind beide Werte häufig positiv oder beide negativ und lassen damit einen gleichgerichteten Zusammenhang zwischen ihnen vermuten. Eine Einteilung in vier Quadranten erfolgt entlang der Nulllinien. 17 Fahrer sind in beiden Kriterien besser und haben daher positive Werte für beide Kriterien (Q1). Nur fünf Fahrer waren schneller als die schnellste Route und haben dafür längere Routen gewählt (Q2). Die meisten Fahrer ($N=66$) haben zwei negative Werte und schaffen es damit nicht bei einem der beiden Kriterien einen besseren Wert zu erreichen als die berechnete schnellste Route (Q3). In Q4 befinden sich 18 Fahrer, die kürzere Routen als die schnellste Route auswählen, dafür aber auch mehr Zeit benötigen. Die beiden Maße M_d für die Routendauer (Dimension der Zeit) und M_l für die Routenlänge (Dimension des Raumes) der Routenperspektive gehen in die Modellierung ein (Modell 2a und 2b).

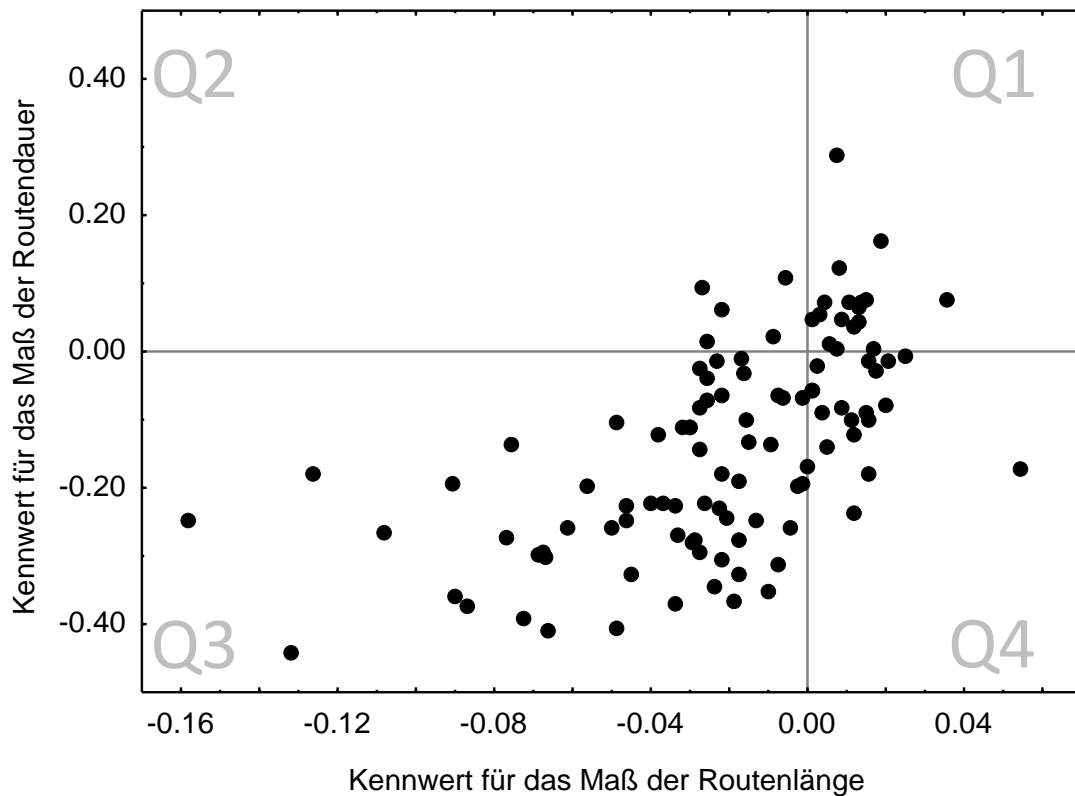


Abbildung 7.8: *Ergebnis der Modellierung der Routenperspektive pro Fahrer*

7.3 Ergebnisse für die Netzwerkperspektive

Für die Analyse des Mobilitätsverhaltens der Fahrer für die Netzwerkperspektive wurden in der Datenaufbereitung alle gefahrenen Kanten bzw. Routen eines Fah-

ers in einen Graphen umgewandelt, siehe Abschnitt 6.3.3. Dieser Graph wird mit Hilfe graphentheoretischer Analysewerkzeuge untersucht, um die zwei Indikatoren für die Größe und die Struktur zu bestimmen. Das entsprechende Verfahren (siehe Abschnitt 5.2.3) zur Ermittlung des darauf aufbauenden Maßes M_a ist komplexer als für die Maße der anderen Perspektiven und bildet den wichtigsten Baustein zur Einbeziehung der Dimension des Raumes in die Modellierung und entsprechend für die Personalisierung.

7.3.1 Indikator der Größe

Wichtigste Kennwerte eines Graphen sind die Anzahl an Kanten und Knoten. In Abbildung 7.9, links sind die Werte für die Fahrer dargestellt. Die Mediane liegen bei 9000 Knoten und 10000 Kanten. Das bedeutet, dass die Adjazenzmatrizen der erstellten Graphen nicht sehr stark besetzt sind, da es nur unwesentlich mehr Kanten gibt. Analog zu den insgesamt gefahrenen Kilometern sind auch die Anzahl an Knoten und Kanten sehr verschieden für die einzelnen Fahrer. Es gibt Fahrer für die die Knoten- bzw. Kantenanzahl bei 3000 liegen und Fahrer mit über 20000 Knoten bzw. Kanten. Die meisten Fahrer bewegen sich allerdings in einem Bereich zwischen 6000 und 11000 Knoten sowie 7000 und 14000 Kanten.

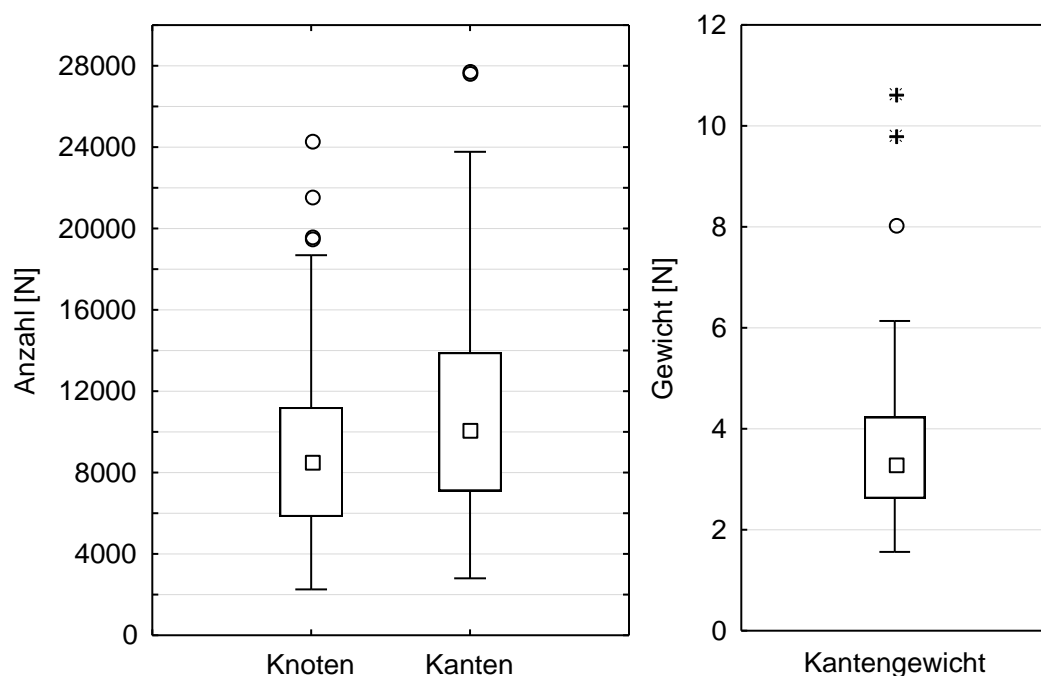


Abbildung 7.9: Verteilung der Knoten- und Kantenanzahl sowie der Kantengewichte

In Abbildung 7.9, rechts sind die Mittelwerte der Kantengewichte pro Fahrer dargestellt. Sie bilden die Grundlage für den dritten Kennwert zur Berechnung des Indikators Größe. Er zeigt an, wie häufig eine Kante bzw. Straße genutzt wurde. Hier liegt der Median bei drei, allerdings gibt es auch drei Fahrer, die viele Kanten häufiger besuchen und Mittelwerte von mehr als acht aufweisen können.

Eine Min-Max-Normierung wird genutzt, um die Werte in Kennwerte K_{pf} , K_{qf} und K_{gf} zu überführen und sie für die Indikatorberechnung nutzbar zu machen. Das Ergebnis für die drei Kennwerte ist in Abbildung 7.10 dargestellt. Für jeden Kennwert gibt es aufgrund der Definition mindestens einen Null- und einen Einswert. Die anderen Fahrer gruppieren sich dazwischen. Für Knoten und Kanten liegt der Median bei 0.3 für die Kantengewichte bei 0.2. Da die entsprechenden Boxen recht ausgeglichen sind, haben mehr als 75% der Fahrer einen Wert kleiner als 0.5. Dies lässt vermuten, dass viele Fahrer zwar ähnlich große Graphen erzeugen, es nach oben aber einige extreme Ausreißer gibt, die das Bild entsprechend auseinanderzerren. Dies müssen nicht unbedingt Fahrer sein, die schon bei der Beschreibung der Stichprobe in Abschnitt 4.2.1 auffällig waren. Durch die Umwandlung in einen Graphen ist die längentechnische Komponente ausgeklammert worden.

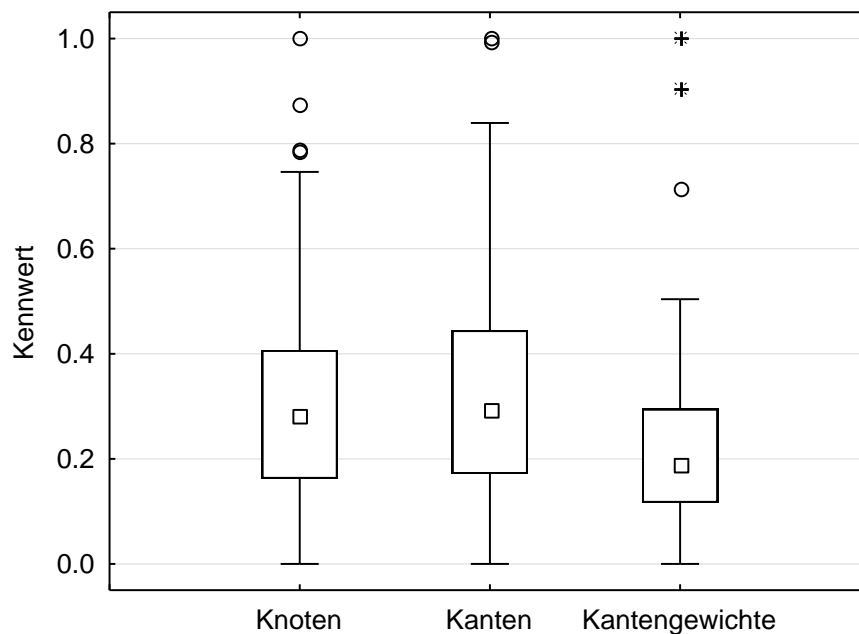


Abbildung 7.10: *Normierte Kennwerte für Knoten, Kanten und Kantengewichte*

Die drei Werte werden miteinander kombiniert, um Indikatorwerte $I_{Größe_f}$ zu erhalten. Dabei gehen sie mit gleichem Gewicht ein, das Ergebnis ist in Abbildung

7.11 für alle Fahrer aufgetragen. Meist bewegen sich Werte für den Indikator Größe zwischen 0.2 und 0.4. Der Werte null oder eins kommen nicht vor. Das bedeutet, dass es keinen Fahrer gibt, der bei allen drei Werten das Maximum bzw. das Minimum erreicht. Der so ermittelte Wert beschreibt die Größe des Graphen, der durch das Mobilitätsverhalten des Fahrers aufgespannt wurde.

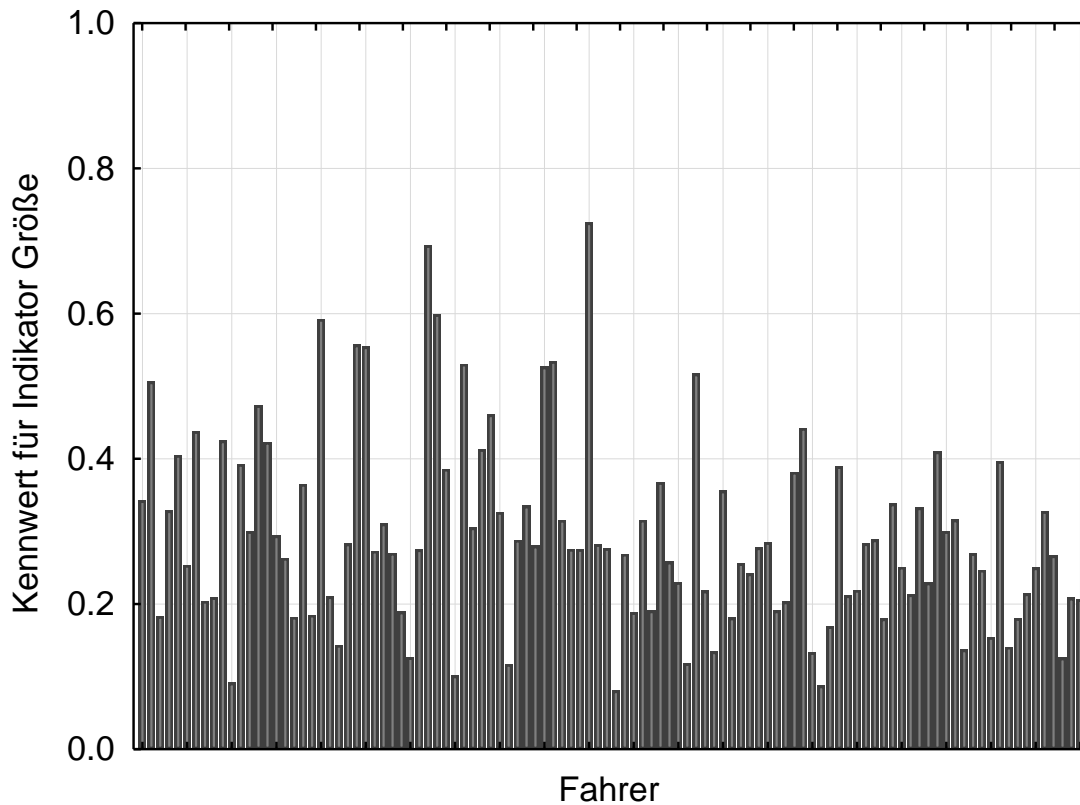


Abbildung 7.11: Wert des Indikators Größe pro Fahrer

7.3.2 Indikator der Struktur

Der Indikator der Struktur umfasst ebenfalls drei Kennwerte. Zum einen ist dies die Anzahl der Kreise, die mit der Eulerschen Polyederformel ermittelt wird. In Abbildung 7.12, links ist die Anzahl der Kreise abgebildet. Der Median liegt bei 1500 Kreisen. Allerdings gibt es auch Fahrer mit weniger als 1000 und andere Fahrer mit mehreren Tausend Kreisen. Die Mehrheit der Fahrer bewegt sich aber im Bereich zwischen 500 und 3000 Kreisen im Graphen.

Zum anderen werden der Kern sowie der innere Kern, der sich aus dem Kantengewichtszusammenhang (siehe Abschnitt 5.2.3) ergibt, ermittelt. Der Kern umfasst die größte Knotenuntermenge, die mit Kanten eines Gewichts von mindestens zwei

verbunden ist (Kern2). Für den inneren Kern wird das Mindestgewicht auf fünf festgelegt (Kern5). Daher ist der innere Kern kleiner als der Kern. Für diesen liegt der Median bei 2400 Knoten, die durch Kanten mit einem Gewicht größer eins verbunden sind. Dieser Kern kann bei einigen Fahrern auch mehr als 5000 Knoten umfassen, für die Hälfte der Fahrer liegt die Anzahl allerdings zwischen 1500 und 3000 Knoten. Der Median des inneren Kerns beträgt 800 Knoten und nur wenige Fahrer haben einen inneren Kern bestehend aus mehr als 2000 Knoten.

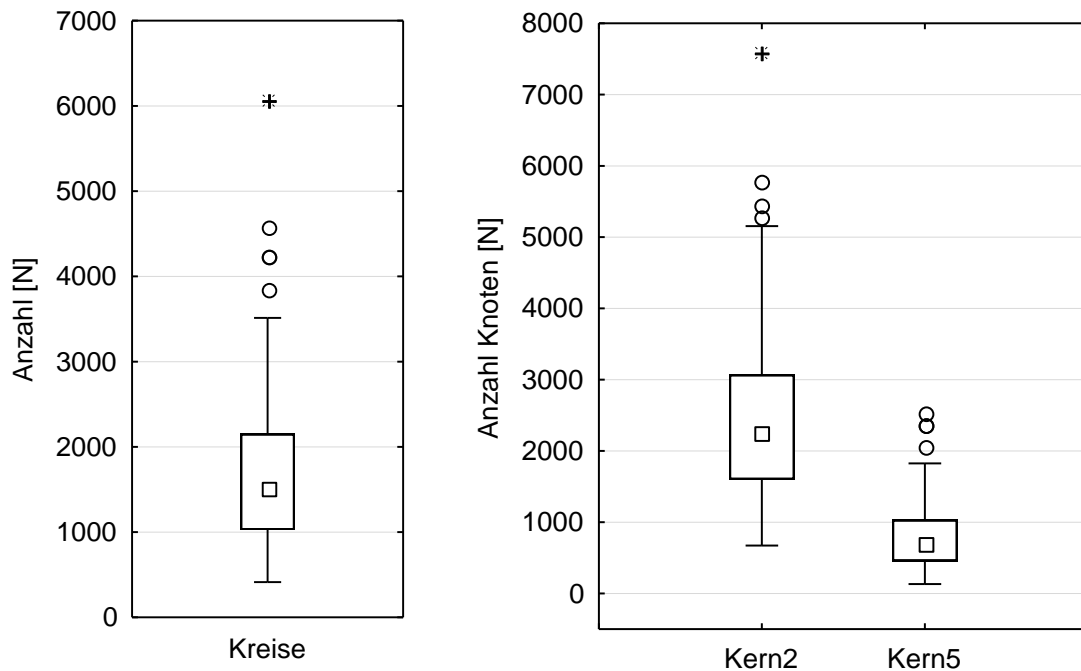


Abbildung 7.12: Verteilung der Anzahl an Kreisen (links) sowie an Knoten in den Kernen (rechts)

Um eine Vergleichbarkeit der Kerne zu erreichen, wird die relative Größe der Kerne im Vergleich zum gesamten Graphen benutzt. Somit werden die verschiedenen Knotenanzahlen miteinander verglichen. Das Ergebnis ist in Abbildung 7.13 für die beiden Kerne aufgetragen. Die Mediane liegen bei 0.3 und 0.08 (Maximum bei 0.6 und 0.3, Minimum bei 0.08 und 0.01). Die Streuung ist ähnlich wie bei den absoluten Zahlen, wobei die Mediane in einem niedrigeren Bereich als im Vergleich zu den jeweiligen gesamten Verteilungen liegen.

Für alle drei Parameter wird erneut eine Min-Max-Normierung durchgeführt, siehe Abbildung 7.14. Wie schon bei der Berechnung der Kennwerte für den Indikator Größe sind auch hier die erhaltenen Kennwerte (K_{tf} , K_{Kernf} und $K_{inKernf}$) deutlich kleiner als 0.5 mit größeren Ausreißern nach oben. Für die Anzahl der Kreise liegt

der Median bei 0.2. Der Kern2-Median liegt bei 0.45, was darauf hindeutet, dass die Ausreißer nach oben hier nicht so stark ausgeprägt sind.

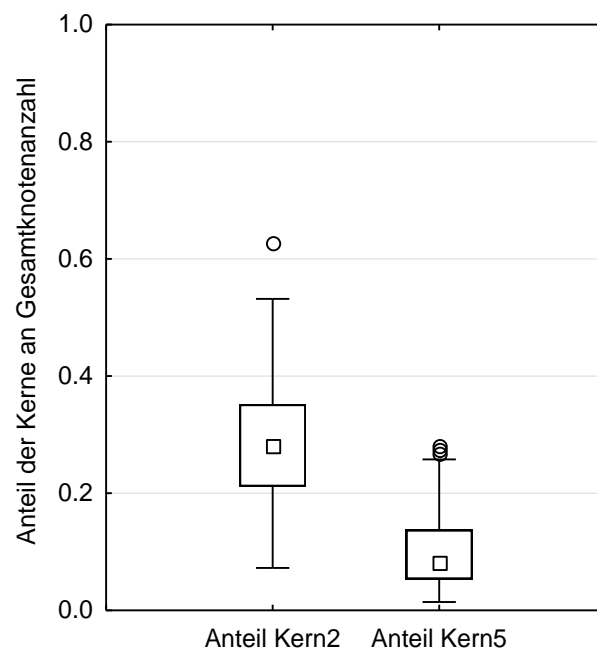


Abbildung 7.13: *Relativer Kennwert der Kerne*

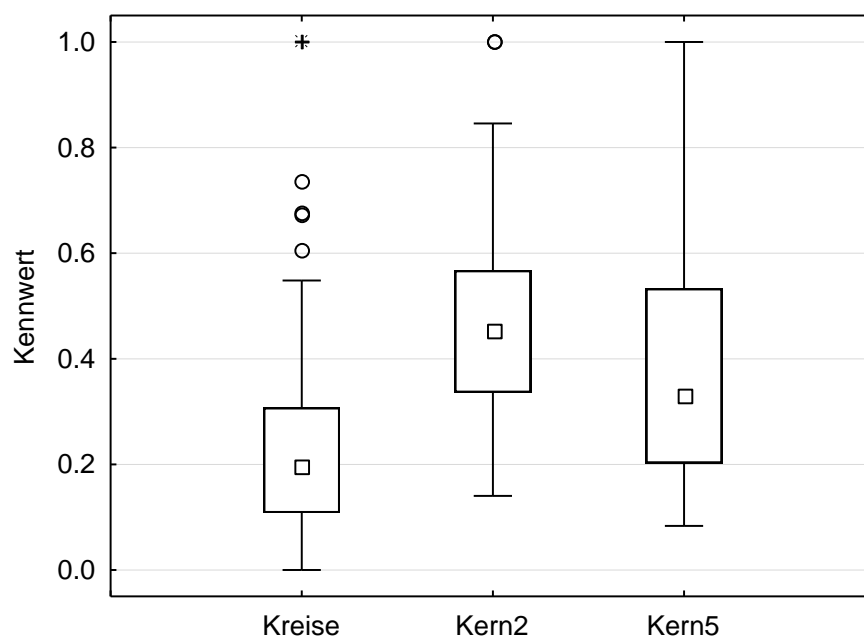


Abbildung 7.14: *Normierte Kennwerte für Kreise sowie der relativen Kerngrößen*

Der Median für den inneren Kern liegt wiederum bei 0.33, wobei anhand der Boxen sowie der Whisker erkennbar ist, dass die Werte eine größere Streuung aufweisen.

In die Berechnung des Indikators Struktur $I_{Struktur_f}$ geht der Parameter für die Kreise mit 0.5 in die Bewertung ein. Die beiden Kerne werden gleich behandelt und haben somit den Faktor 0.25. Das Ergebnis ist in Abbildung 7.15 fahrerweise aufgetragen. Auch hier liegt der Indikatorwert für die meisten Fahrer zwischen 0.2 und 0.4 mit Ausreißern nach oben. Ebenfalls gibt es keinen Fahrer, der für alle drei Parameter das Maximum erreicht hat.

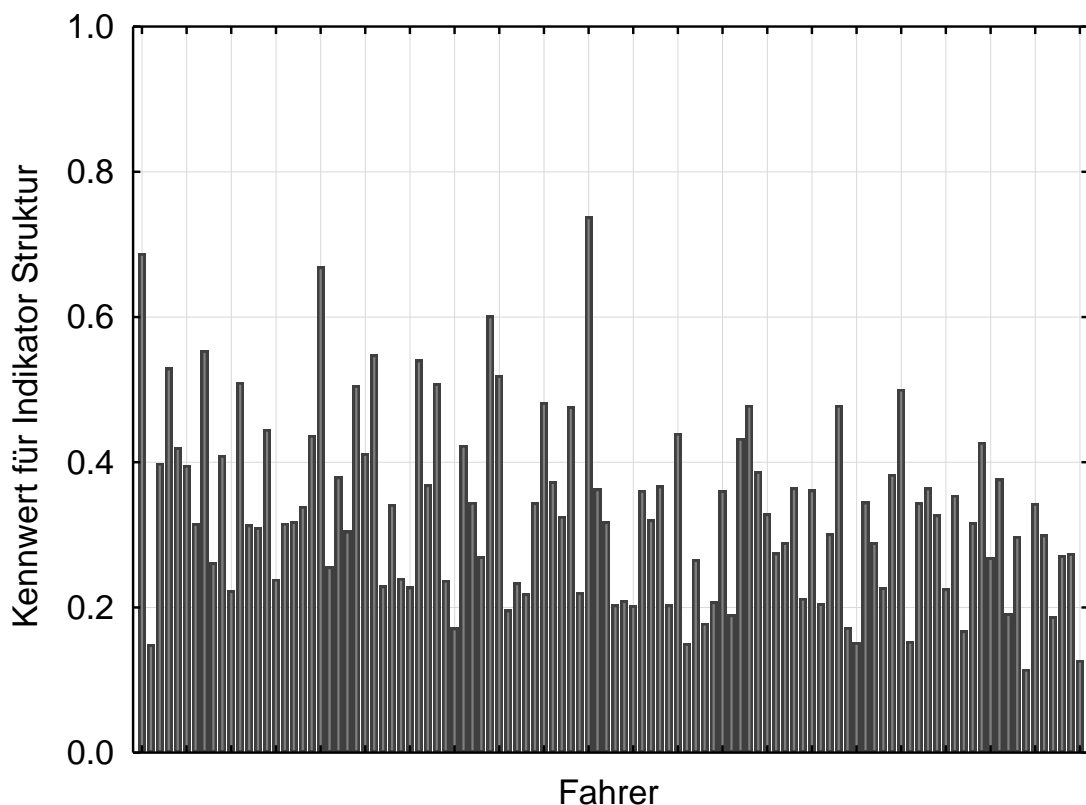


Abbildung 7.15: Wert des Indikators Struktur pro Fahrer

7.3.3 Bestimmung eines Navigationstyps für die Modellerstellung

Die Zuordnung eines Fahrers zu einem Navigationstyp geschieht in Abhängigkeit beider Indikatoren. Eine Visualisierung ist in Abbildung 7.16 aufgetragen. Eine Unterteilung in vier verschiedene Quadranten bzw. Navigationstypen geschieht anhand der berechneten Mittelwerte. Diese sind als Linien in der Abbildung erkenn-

bar. Wie bereits aus den Ergebnissen aus Abbildung 7.11 und 7.15 zu vermuten war, befinden sich die meisten Fahrer in Q3.

28 Fahrer mit großen Werten für beide Indikatoren haben einen großen, netzförmigen Graphen und werden als explorativ klassifiziert (Q1). Dagegen bauen 42 Fahrer mit zwei kleinen Werten einen eher sternförmigen Graphen mit geringer Ausdehnung auf und sind somit festgelegte Fahrer (Q3). Der Graph von 15 Fahrern in Q4 hat zwar eine große Ausdehnung, ist dafür aber eher sternförmig, womit diese den explorativ festgelegten Fahrern zugeordnet sind. 21 Fahrer haben einen Graphen mit netzförmiger Struktur aber nur geringer Ausdehnung generiert (Q2). Diese Fahrer sind lokal explorativ.

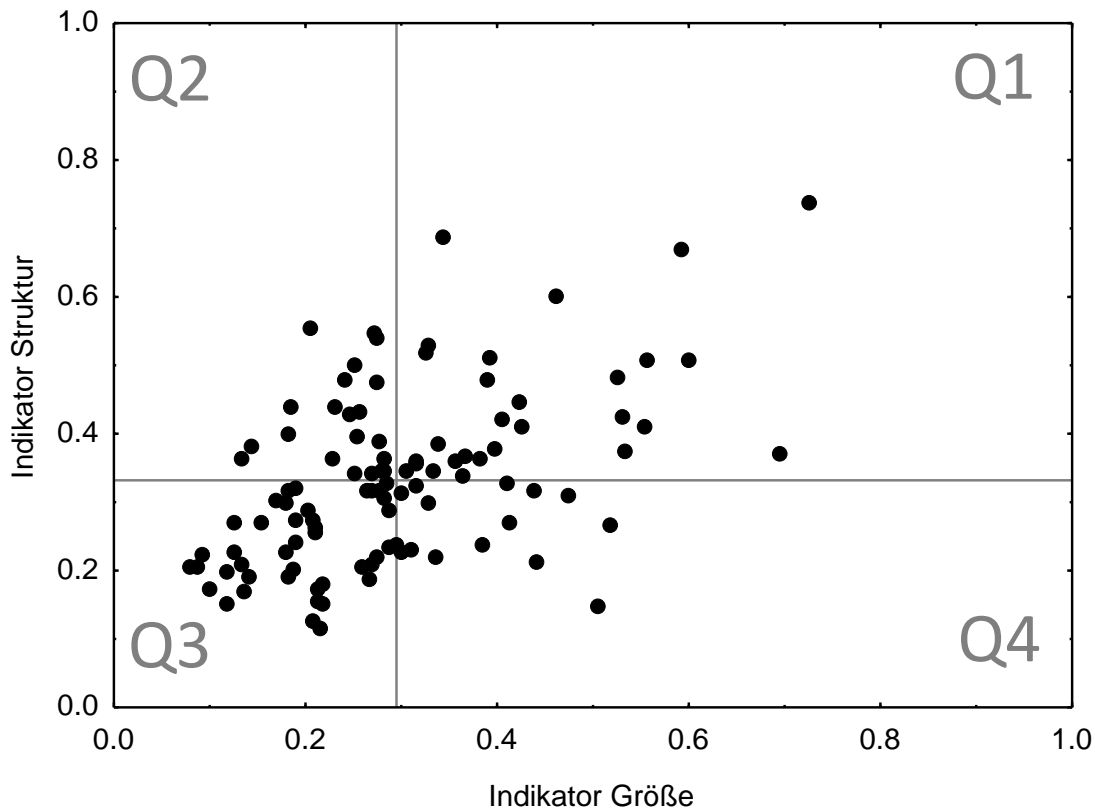


Abbildung 7.16: *Visualisierung des Navigationstyps*

Für die Berechnung eines Kennwerts für die Abdeckung M_a für die Netzwerkperspektive, die der Dimension des Raumes zugeordnet ist, werden die beiden Werte der Indikatoren miteinander kombiniert. Da beide Werte signifikant miteinander korrelieren ($r = 0.5064$) geschieht dies durch Addition und Normierung. Das Ergebnis der Berechnung ist für jeden Fahrer analog zu den vorherigen Abbildungen in Abbildung 7.17 dargestellt. Auch hier sind die Kennwerte für M_{a_f} für die meis-

ten Fahrer niedriger als 0.5, Unterschiede zwischen den Fahrern sind aber deutlich erkennbar. Auch diese Kennwerte werden für die Modellierung im nächsten Abschnitt genutzt.

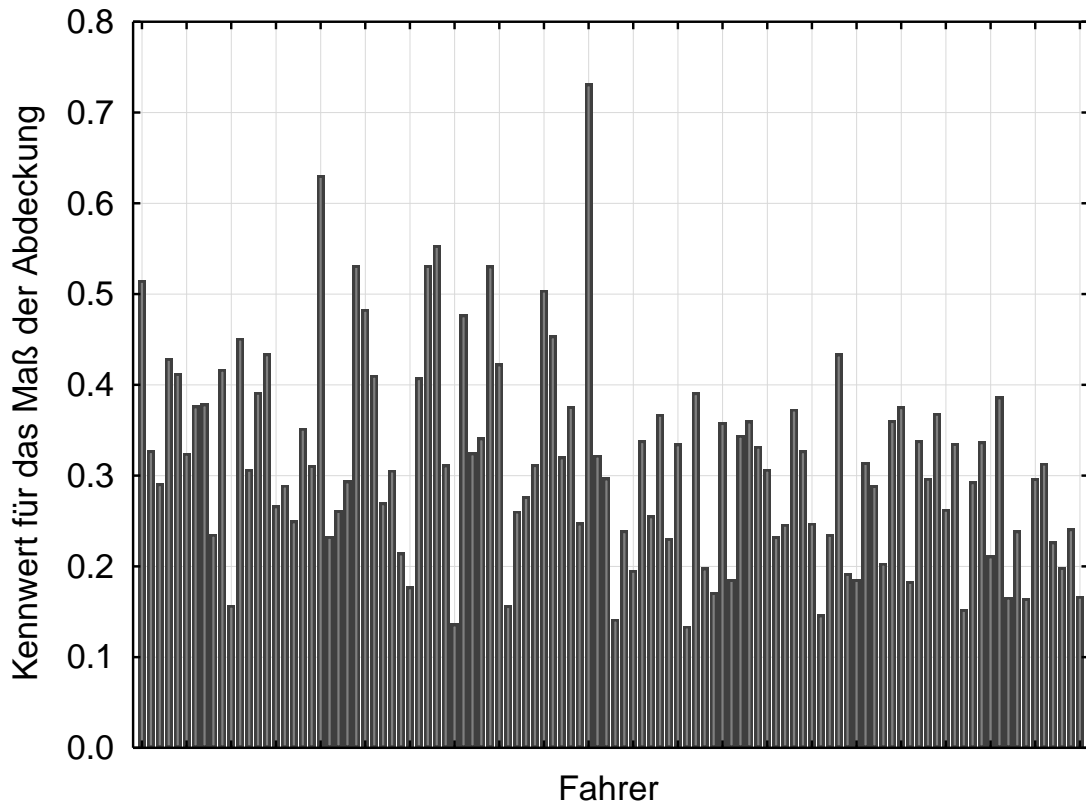


Abbildung 7.17: *Ergebnis der Modellierung der Netzwerkperspektive pro Fahrer*

7.4 Erstellung eines zweidimensionalen Modells

Die Modellierung baut auf den im vorherigen Abschnitt ermittelten Maßen auf. Zunächst wird der Zusammenhang zwischen den Maßen der verschiedenen Perspektiven analysiert, wie in Abschnitt 5.3.1 beschrieben. Anschließend werden jeweils zwei Maße für die Dimension der Zeit sowie für die Dimension des Raumes einbezogen und darauf aufbauend zwei Maße M_z und M_r berechnet.

7.4.1 Durchführung der Korrelationsanalyse

Der Zusammenhang zwischen den Maßen wird sowohl grafisch wie auch mittels statistischer Korrelationsanalyse untersucht. Zunächst werden die Maße gleicher Dimension untersucht, deren Zusammenhang in Abbildung 7.18 illustriert ist. Für

die Dimension der Zeit werden links die Kennwerte des Maßes M_v aus der Kantenperspektive mit den Kennwerten des Maßes M_d der Routenperspektive verglichen. Erneut stellt ein Fahrer einen Punkt dar. Es ist eine sehr große Streuung zwischen den Fahrern zu beobachten. Dazu aber auch eine leichte Korrelation, wie es bei Kennwerten, die die gleiche Dimension betreffen, zu erwarten ist. Ein Fahrer, der schneller fährt, sollte generell auch schnellere Routen wählen.

Rechts werden die Kennwerte des Maßes M_l der Routenperspektive und die Kennwerte des Maßes M_a der Netzwerkperspektive für die Dimension des Raumes betrachtet. Auch hier ist neben einer großen Streuung eine leichte Abhängigkeit erkennbar. Fahrer, die ein großes Netz aufbauen, sollten generell auch aufgrund ihrer Kenntnis des Netzes in der Lage sein, Routen zu wählen, die kürzer sind.

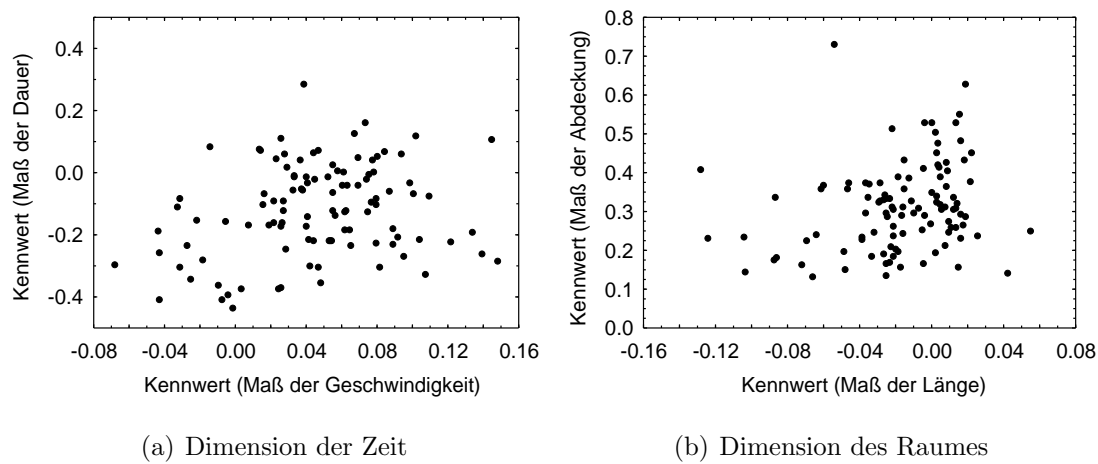


Abbildung 7.18: *Vergleich der Kennwerte gleicher Dimension*

Als nächstes werden Maße unterschiedlicher Dimension verglichen. Dies sind die Kombinationen der Maße M_d und M_l der Routenperspektive sowie die Maße M_v und M_a . Die Maße der gleichen Perspektive zeigen einen stärkeren positiven Zusammenhang als die Maße gleicher Dimension (Abbildung 7.19, links). Dagegen zeigt sich für M_v und M_a kein Zusammenhang (Abbildung 7.19, rechts).

Vielmehr ist bei weiterhin sehr großer Streuung nicht erkennbar, ob ein großer Wert bezüglich der Dimension der Zeit auch einen großen Wert bezüglich des Raumes nach sich zieht. Somit kann nicht direkt gefolgert werden, dass ein schneller Fahrer auch räumlich ein besonders gutes Verhalten zeigt.

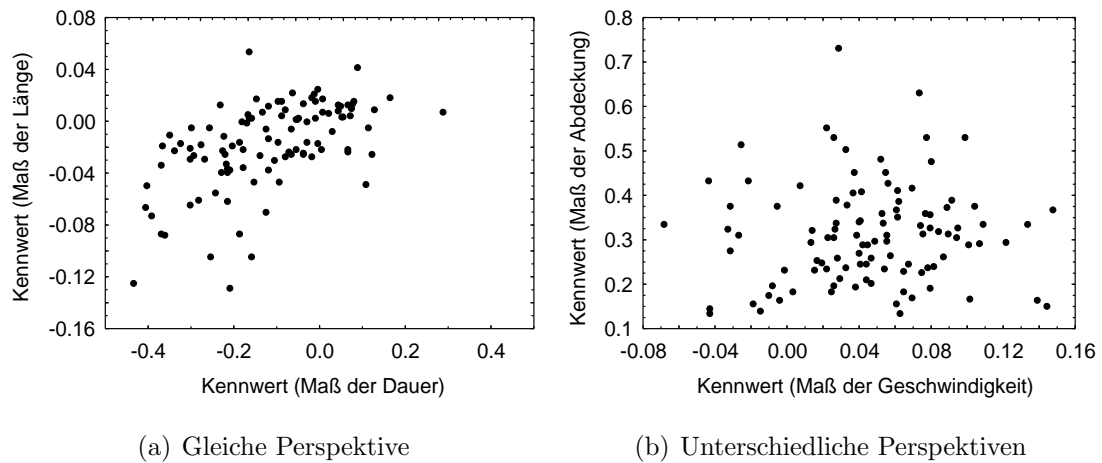


Abbildung 7.19: Vergleich der Kennwerte unterschiedlicher Dimension

Zusätzlich zur grafischen Analyse werden die Zusammenhänge statistisch untersucht. Hierfür wird eine Korrelationsmatrix erstellt, die die Korrelation der verschiedenen Maße mit Hilfe der Pearson-r-Korrelation untersucht. Die Voraussetzung, dass die Werte normalverteilt sind, kann bei der Prüfung mit Hilfe verschiedener Verfahren (Kolmogorov-Smirnov, Lilliefors, Chi-Quadrat [35]) nicht verworfen werden und ermöglicht eine Untersuchung. Die Ergebnisse sind in Tabelle 7.1 dargestellt. Signifikante Ergebnisse sind fett markiert. Zwischen den Maßen M_d und M_l der gleichen Perspektive ist die größte signifikante Korrelation feststellbar. Dazu gibt es nicht sehr stark ausgeprägte aber signifikante Korrelationen zwischen den Kennwerten gleicher Dimension M_v und M_d sowie M_l und M_a .

Maße	M_l	M_d	M_v
M_a	0.2243	0.1788	0.0491
M_l		0.5493	-0.0189
M_d			0.2449

Tabelle 7.1: Korrelationen zwischen den Kennwerten der verschiedenen Perspektiven und Dimensionen

Ebenso wird der Eindruck aus Abbildung 7.19, rechts bestätigt, dass zwischen den Maßen M_v und M_a keine signifikante Korrelation besteht. Damit stützen sich zwar die Maße gleicher Dimension und gleicher Perspektive, doch gibt es keinen direkten Zusammenhang zwischen Maßen verschiedener Dimension. Somit müssen beide Dimensionen für ein umfassendes und fundiertes Modell basierend auf den Maßen der verschiedenen Perspektiven mit einbezogen werden.

7.4.2 Ergebnis der zweidimensionalen Modellierung

Für jeden Fahrer wurden jeweils zwei Kennwerte für Maße für die Dimension Zeit wie auch für die Dimension Raum berechnet. Diese werden jeweils zu einem Wert (M_z bzw. M_r) aggregiert und ermöglichen eine Klassifizierung der Fahrer in Anlehnung an die Abbildung 3.7 aus Abschnitt 3.2.2. Zunächst werden die Fahrer in einem ersten Schritt anhand der Mittelwerte der gesamten Stichprobe unterteilt, siehe Abbildung 7.20.

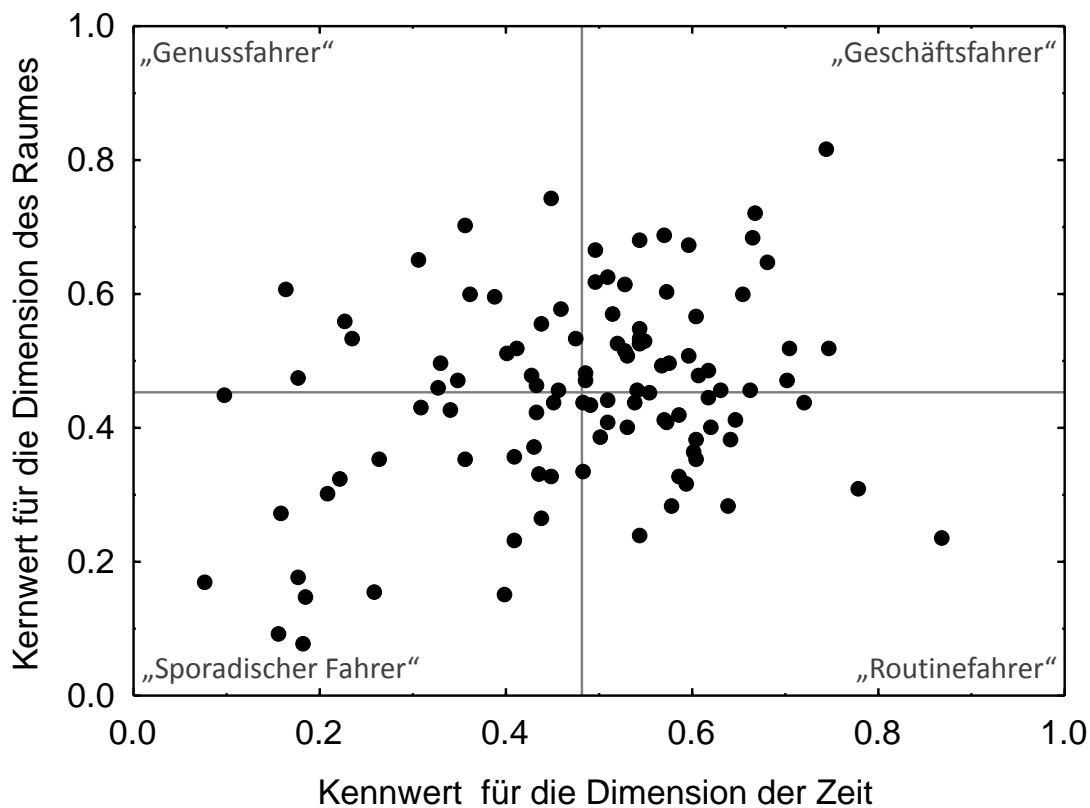


Abbildung 7.20: Mittelwertbasierte Unterteilung der Fahrer in vier Klassen

Fahrer mit zwei kleinen Werten werden als sporadische Fahrer klassifiziert, die langsam und wenig fahren und dabei die Routen nur selten variieren. Diese Klasse umfasst 25 Fahrer. Dagegen steht die Klasse der Geschäftsfahrer mit 35 Fahrern. Diese Fahrer sind schnell und viel unterwegs. Dabei nutzen sie ihre Kenntnisse, um die Strecken zu variieren und Abkürzungen zu nehmen, um so einen großen Wert bezüglich des Raumes zu erreichen.

Schnell fahren ebenfalls die Routinefahrer, allerdings haben sie eher feste Strecken mit einer geringeren Reichweite und Variabilität und damit einen niedrigen Wert für die Dimension des Raumes. Diese sind in der Einordnung rechts unten mit

26 Fahrern zu erkennen. Demgegenüber steht die Klasse der Genussfahrer mit 20 Fahrern. Diese fahren eher langsam, dafür aber ohne festgelegtes Muster und damit mit größerer Variabilität.

Eine Schwäche der Einteilung anhand des Mittelwerts ist die Zuordnung von Fahrern in der Mitte der Stichprobe mit sehr ähnlichen Merkmalen aber unterschiedlichen Klassifizierungen, da diese nahe an der Entscheidungsgrenze liegen. Dies kann zu einem Auseinanderreißen von Fahrern mit ähnlichen Merkmalen führen. Daher wird im Folgenden ein k-means-Clusterverfahren [12] mit $k=4$ angewandt, um vier Gruppen zu bestimmen, siehe Abbildung 7.21.

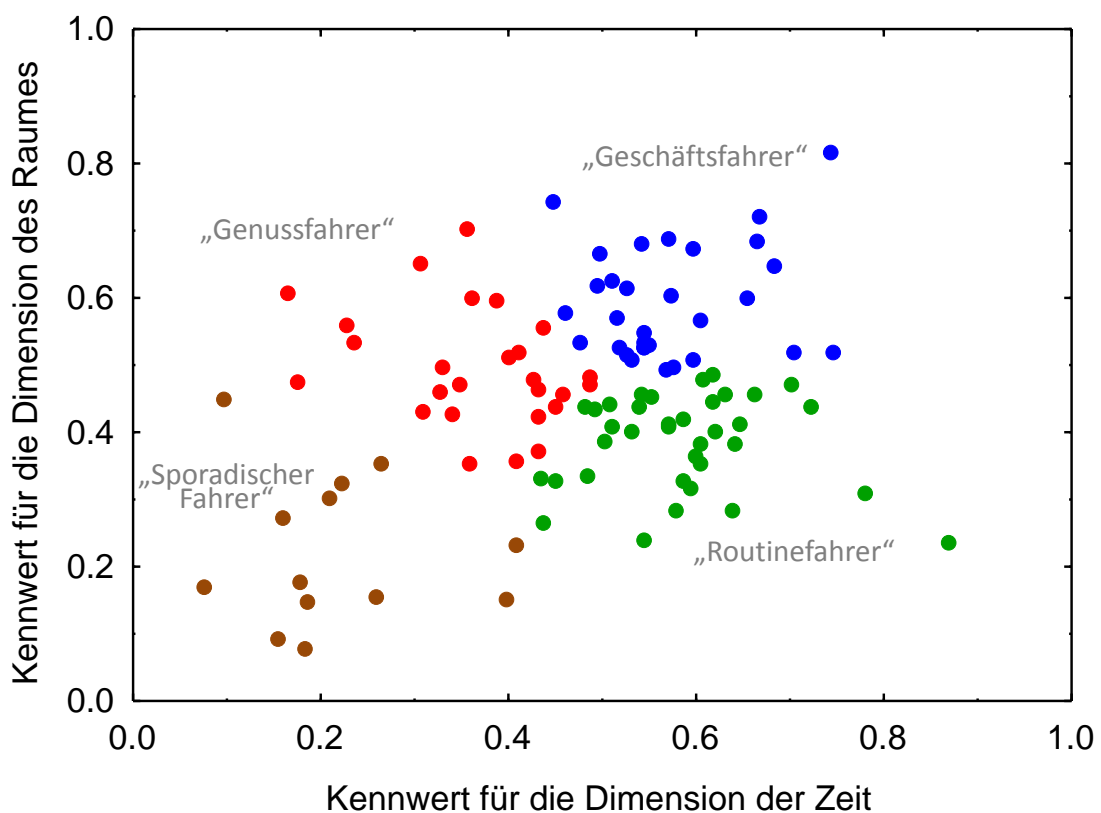


Abbildung 7.21: Klassifizierung der Fahrer mit Hilfe eines Clusterverfahrens

Die Fahrer an den Rändern der Stichprobe sind überwiegend in der gleichen Klasse geblieben. Die größten Veränderungen gegenüber der Mittelwerteinteilung fanden im Zentrum statt. Die Gruppe der Geschäftsfahrer hat sich durch das Clusterverfahren nur unwesentlich verändert und besteht aus 31 Fahrern. Dahingegen werden nur noch 13 Fahrer als sporadisch klassifiziert.

Die Klassen der Genussfahrer und Routinefahrer bestehen nun aus 26 bzw. 36 Fahrern und beinhalten auch die Fahrer mit mittleren Kennwerten. Insgesamt ist eine

auf Clusterverfahren basierende Klassifizierung zur deskriptiven Beschreibung der Fahrer und ihres navigatorischen Verhaltens sinnvoll. Eine Klassifizierung in mehr oder weniger Gruppen ist zwar möglich, bringt aber keinen weiteren Erkenntnisgewinn für die folgende Personalisierung. Die ermittelten Gruppen werden bei der Bewertung der Personalisierung in Abschnitt 7.6 genutzt.

7.5 Umsetzung der Personalisierung

Aufbauend auf den Ergebnissen der Modellierung erfolgt die Personalisierung. Dabei werden drei Ansätze umgesetzt. Im ersten Schritt erfolgt die Personalisierung der Kantengewichte. Der zweite Schritt umfasst die Personalisierung der Zielfunktion und im dritten Schritt wird eine kombinierte Personalisierung durchgeführt. Das Vorgehen folgt der Beschreibung in Abschnitt 5.4.

7.5.1 Personalisierung der Kantengewichte

Die Anpassung der Kantengewichte basiert direkt auf den Ergebnissen der Kantenspektivanalysen. Im Gegensatz zur Modellierung (siehe Abbildung 7.4) werden hierfür die direkten relativen Abweichungen auf Basis der in der digitalen Karte hinterlegten Geschwindigkeiten bestimmt, siehe Abbildung 7.22. Dazu ist in Abbildung 7.23 ein Ausschnitt der Ergebnisse mit den Geschwindigkeitsklassen ab 40 km/h aufgetragen.

Sehr große Abweichungen gibt es bei der Klasse 10 km/h. Der Median liegt bei 220%, doch gibt es auch Fahrer mit mehr als 400% Abweichung. Diese Kanten sind meist Feldwege oder sehr kleine, enge Straßen, für die nur eine niedrige Geschwindigkeit von 10 km/h angenommen, von den Fahrern aber deutlich höhere Geschwindigkeiten erreicht werden. Dies kann bei der Routenberechnung berücksichtigt werden. Hierbei ist allerdings anzumerken, dass es wiederum von verschiedenen Interessengruppen (Verkehrszentrale, Anwohner,...) auch nicht erwünscht ist, diese Kanten aufzuwerten und somit ein erhöhtes Verkehrsaufkommen zu generieren.

Ähnlich wie bei der Modellierung liegen die Werte für die Klassen größer 20 km/h näher bei denen in der digitalen Karte gespeicherten als für die niedrigeren Klassen. Dementsprechend ist auch die Abweichung geringer, wenn auch vorhanden, siehe Abbildung 7.23. Für die Klasse 20 km/h ist die Abweichung mit 50% im Median deutlich größer.

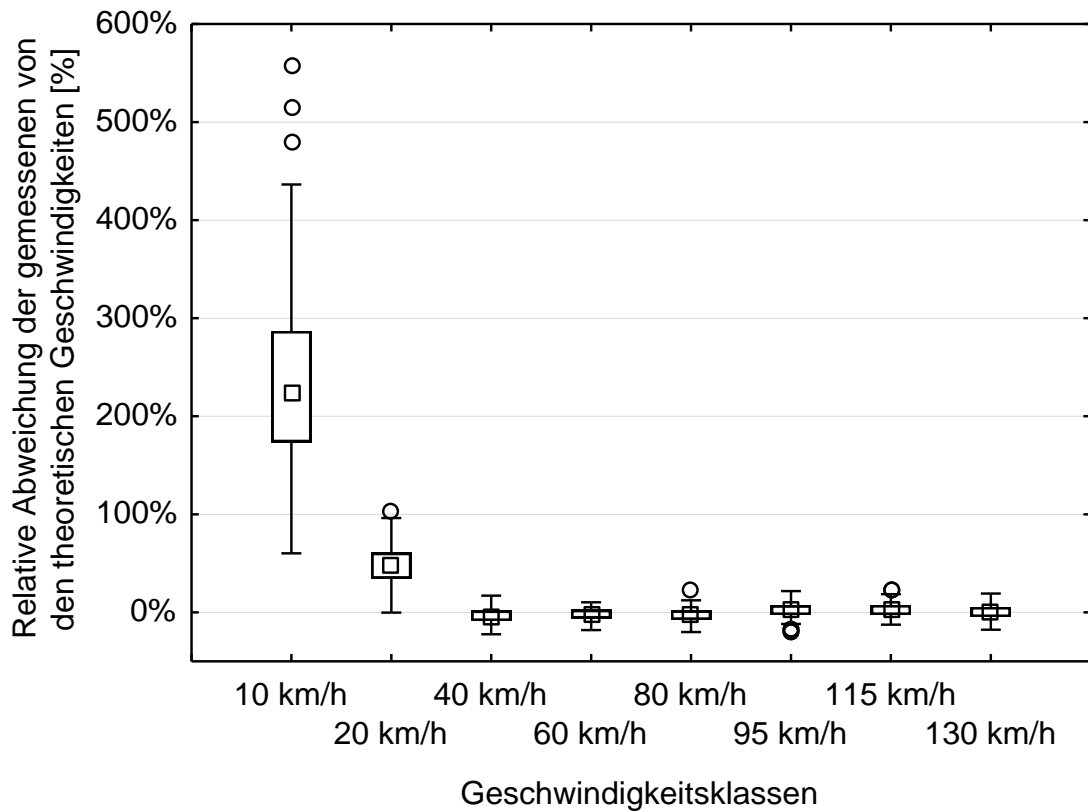


Abbildung 7.22: Abweichung der gemessenen von denen in der digitalen Karte gespeicherten Geschwindigkeiten

Das Ergebnis veranschaulicht ebenfalls die Nutzung des 0-Relativ-Vergleichs für die Modellerstellung. Die Abweichungen für die niedrigen Klassen sind zum Teil sehr groß und würden die Modellierung bei Einbeziehung dominieren, was zu unerwünschten Effekten führen könnte.

Die neu berechneten Kantengewichte bilden die Grundlage für die Personalisierung (PG-Routen) der Routenberechnung. Die Neuberechnung lässt sich wenig rechen- und speicherintensiv durchführen und wird mit Hilfe der internen Routingreferenz durchgeführt.

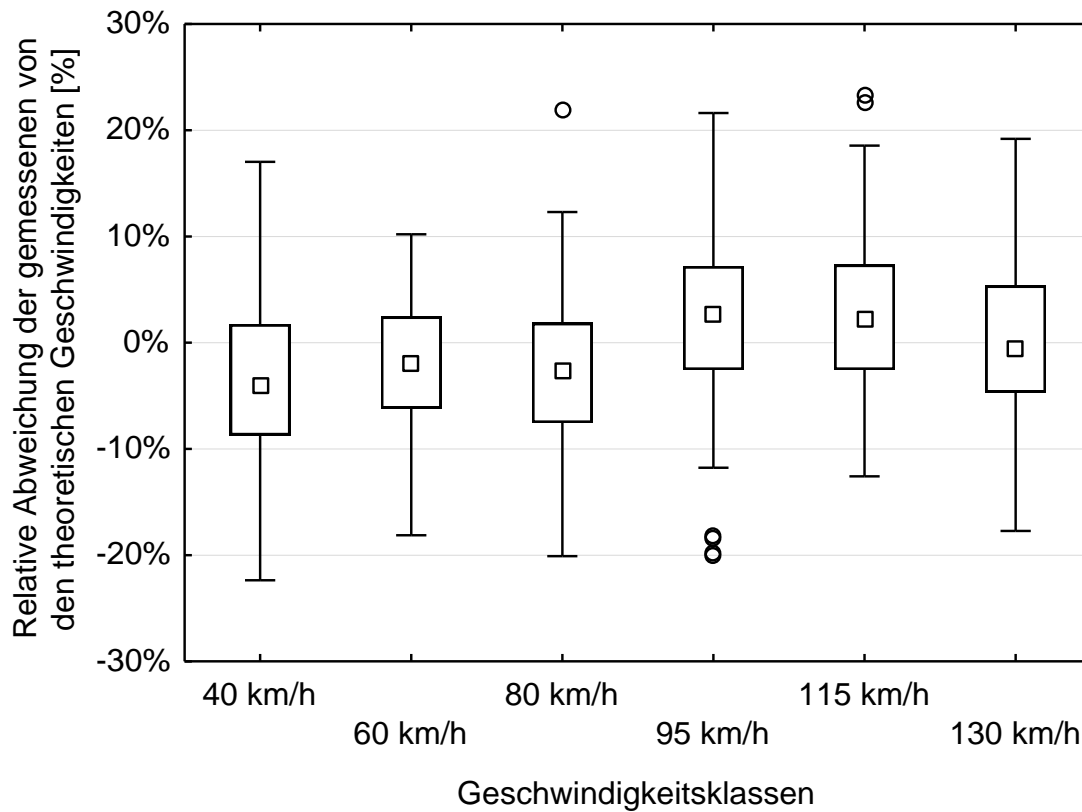


Abbildung 7.23: Abweichung der Geschwindigkeit für die verschiedenen Geschwindigkeitsklassen, Ausschnitt

7.5.2 Personalisierung der Zielfunktion

Der zweite Ansatz beinhaltet die Anpassung der Zielfunktion und basiert auf den beiden normierten Werten für die Maße $M_z(norm)$ und $M_r(norm)$. Nach der Normierung gehen diese als Anteile in die Zielfunktion ein, siehe Abbildung 7.24. Für viele Fahrer liegt das Verhältnis zwischen 40% und 60%. Die Extremwerte liegen in beide Richtungen bei einem Verhältnis von 20% zu 80%. Generell gibt es eine leichte Tendenz zur Dimension der Zeit. Auch für diese personalisierte Zielfunktion wird eine personalisierte Routenberechnung durchgeführt (PZ-Routen).

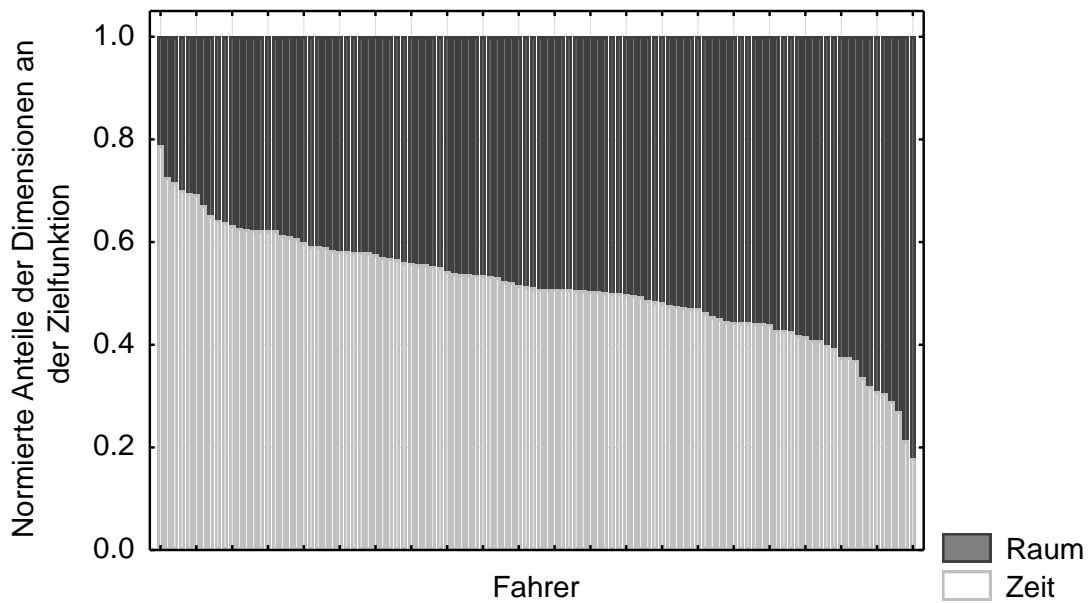


Abbildung 7.24: *Anteile der zweidimensionalen Zielfunktion, sortierte Darstellung*

7.5.3 Kombinierte Personalisierung

In einem letzten Schritt werden die beiden vorherigen Ansätze miteinander kombiniert und eine Routenberechnung mit angepasster Zielfunktion auf Basis neu berechneter Kantengewichte durchgeführt (PK-Routen). Die Anzahl der Individualrouten für alle drei Ansätze, wie auch für die ST-Routen, die zur Bewertung herangezogen werden, sind in Abbildung 7.25 fahrerweise aufgetragen. Wie bereits beschrieben, werden nur FR-Routen mit einer Mindestlänge von einem Kilometer für die Routenberechnung genutzt. Dies sind mehr als 24000 Routen und deckt sich mit den berechneten Routen der Routenperspektive, wobei nun alle Routen berechnet werden können, da die digitale Karte der internen Routingreferenz auf den gemessenen Routen basiert. Somit existiert für jede Start-Zielkombination mindestens eine zulässige Lösung.

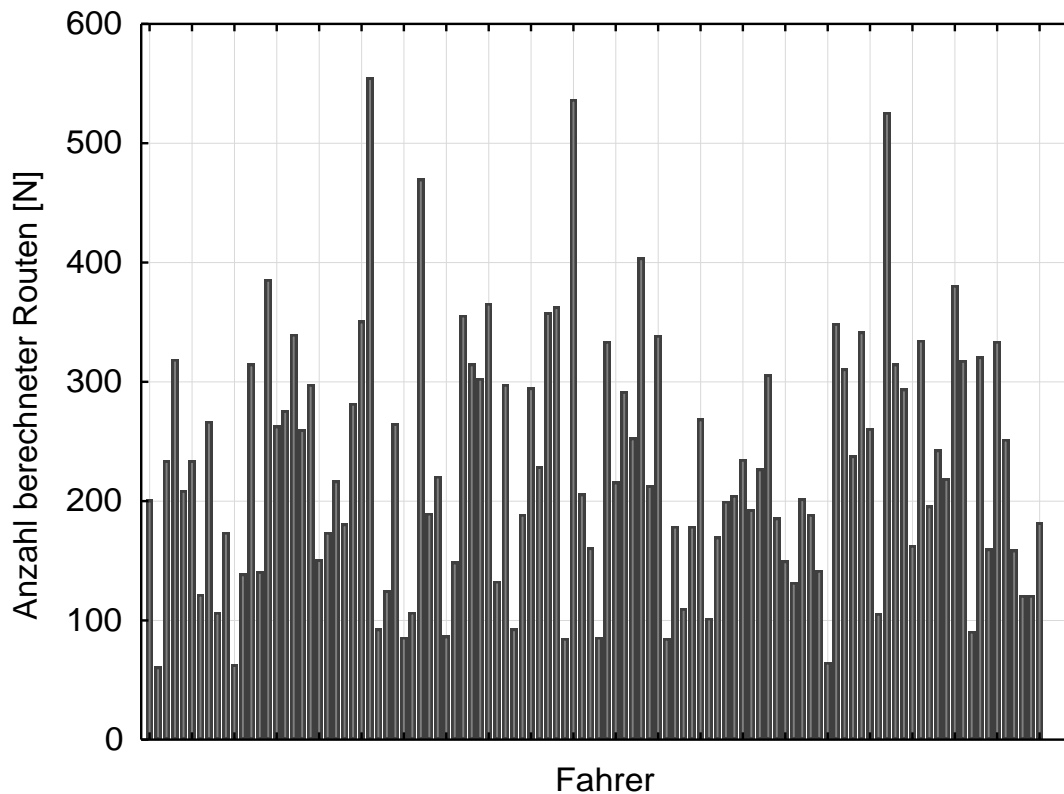


Abbildung 7.25: *Anzahl berechneter Routen pro Fahrer*

7.6 Ergebnisse der Bewertung der Personalisierung

Nach der Umsetzung der Personalisierung werden die Ansätze bewertet. Hierfür werden verschiedene Gütemaße genutzt, um so die Veränderung zu quantifizieren, siehe Abschnitt 5.5. Das erste Gütemaß betrifft die Erreichung der Zielkriterien für die Routendauer und die Routenlänge. Zur Einordnung des Maßes werden zunächst die berechneten mit der gemessenen Route verglichen.

Anschließend erfolgt der Vergleich der berechneten Routen untereinander auf Basis der Zielkriterien. Zusätzlich wird auch der Einfluss der Dimension des Raumes explizit analysiert. Dazu werden die alternativen Gütekriterien der Streckencharakteristika und der Übereinstimmung der berechneten Routen mit der gemessenen Route genutzt und die Ergebnisse diskutiert. Alle Ergebnisse werden in Abhängigkeit vom Routingtyp dargestellt

7.6.1 Nutzung der Zielkriterien als Gütemaß

Vergleich der berechneten mit den gemessenen Routen

Im ersten Schritt werden die berechneten Routen eingeordnet, indem die Zielkriterien mit denen der FR-Routen verglichen werden. Dafür wird für jeden Kantenzug einer gemessenen Route theoretische Werte für die Routendauer und Routenlänge berechnet. Die vier berechneten Routen werden sowohl für die Dimension der Zeit wie auch für die Dimension des Raumes in Relation zu diesen theoretischen Werten gesetzt. Für die Dimension der Zeit sind die Ergebnisse der vier Routen jeweils für die vier Gruppen bzw. Klassen der Routingtypen in Abbildung 7.26 dargestellt.

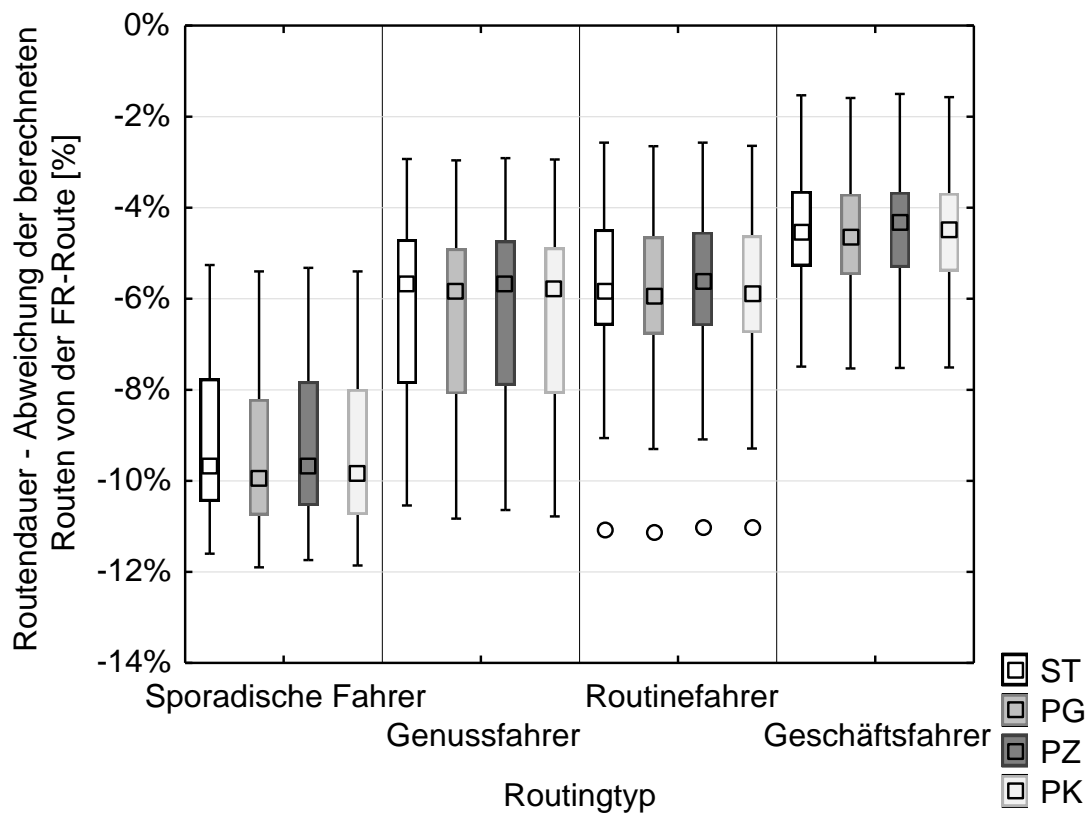


Abbildung 7.26: Abweichung der Routendauer

Bei generell sehr hoher Streuung sind die verschiedenen Klassen anhand der Abweichung erkennbar. Für sporadische Fahrer ist die Abweichung generell größer als für die anderen Klassen. Der Median liegt für alle berechneten Routen bei -10%. Demgegenüber haben Geschäftsfahrer generell eine geringere Abweichung. Hier liegt der Median zwischen -4 und -5%. Dazwischen liegen die Genuss- sowie die Routinefahrer mit einem Median von ungefähr -6%. Damit unterstützen die stabilen Ergebnisse für alle berechneten Routen die Modellierung und die dazugehörige

Klassifizierung. Die Routendauer innerhalb einer Gruppe hingegen verändert sich nur unwesentlich durch die Personalisierung. Eventuell vorhandene Auswirkungen können auch aufgrund der großen Streuung nicht quantifiziert werden.

Bei der Analyse der Routenlänge zeigt sich ein ähnliches Bild für die verschiedenen Klassen, siehe Abbildung 7.27. Sporadische Fahrer haben generell die größte und Geschäftsfahrer die kleinste Abweichung. Damit unterstützen auch die Ergebnisse für die Routenlänge die Modellierung. Dazu zeigt sich innerhalb einer Klasse eine abnehmende Tendenz beim Median. Das bedeutet, dass sich mit zunehmendem Personalisierungsgrad die Routenlänge verkürzt und lässt vermuten, dass die Personalisierung zu einer Verbesserung führt.

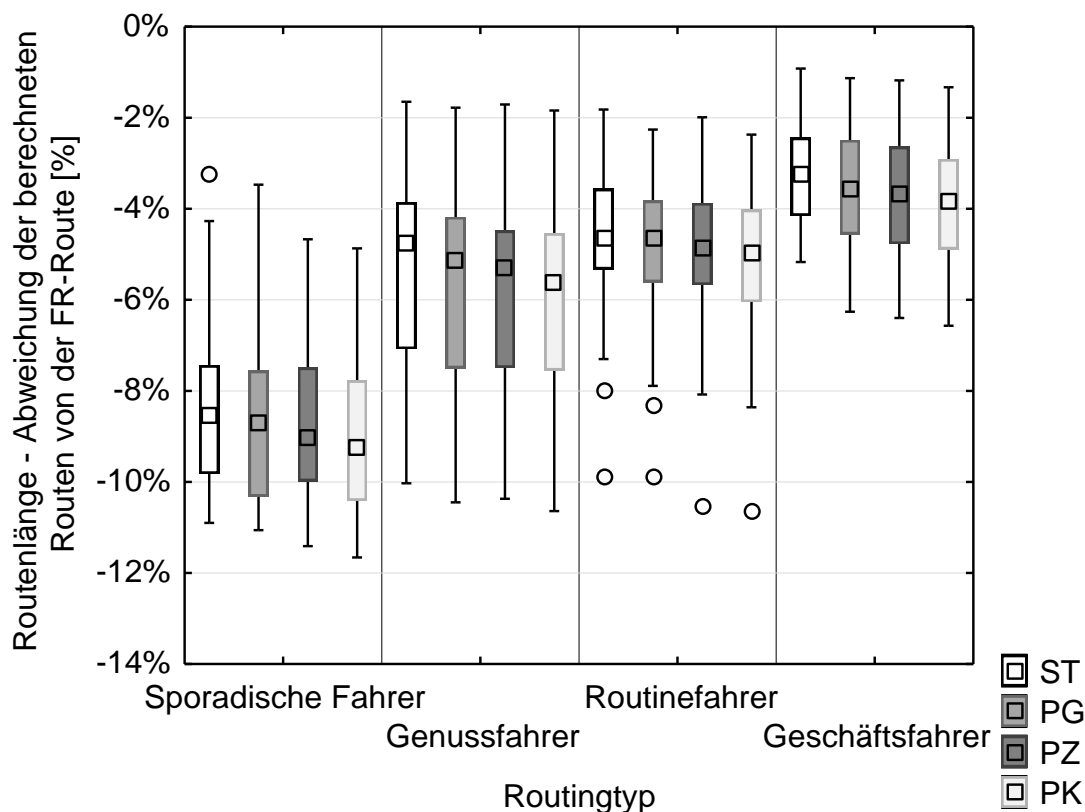


Abbildung 7.27: Abweichung der Routenlänge

Bewertung der verschiedenen Personalisierungsansätze

Um eine potentielle Verbesserung durch die Personalisierung zu quantifizieren, werden die berechneten Routen untereinander bezüglich der Zielkriterien verglichen. Die personalisierte Route mit ihrem Höchstmaß an genutzten Informationen über den Fahrer bildet die Referenz. In Bezug zu dieser Route werden die Abweichungen der ST-, PG- und PZ-Route ermittelt. Dabei beziehen sich die Angaben zur

Veränderung weiterhin auf den Unterschied zur gemessenen Route. In Abbildung 7.28 sind die Ergebnisse erneut nach den verschiedenen Klassen gruppiert dargestellt.

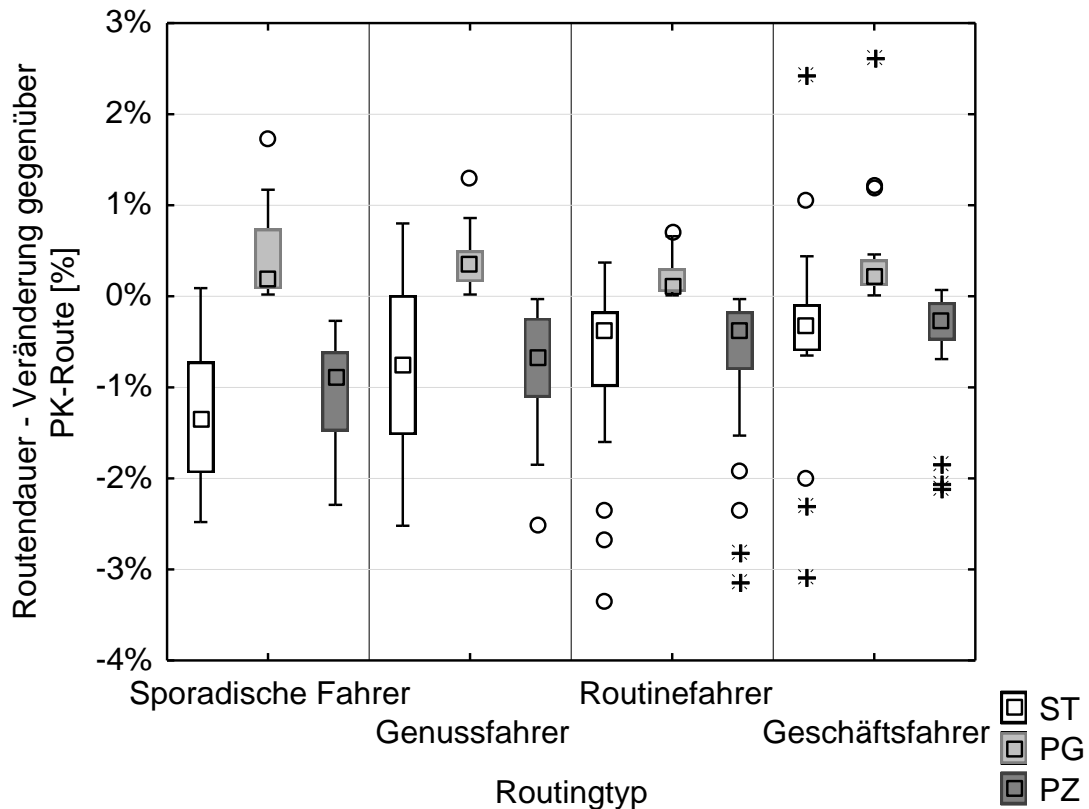


Abbildung 7.28: Bewertung der Routendauerabweichung

Im Vergleich zur kombinierten Personalisierung hat die Nichtberücksichtigung der personalisierten Kantengewichte im Fall der PZ-Routen eine Verschlechterung für nahezu alle Fahrer zur Folge. Besonders ausgeprägt ist die Verschlechterung für sporadische aber auch für Genussfahrer mit einem Median von -0.5 bis -1%. Diese Fahrer haben nur eine geringe Präferenz für die Dimension der Zeit. Dementsprechend sollten die personalisierten Kantengewichte eher zu einer Reduzierung der angenommenen Geschwindigkeit führen. Eine Nichtberücksichtigung sollte daher eine Verschlechterung zur Folge haben. Bei den anderen beiden Gruppen mit einer hohen Präferenz für die Zeit ist der Effekt nicht so stark ausgeprägt.

Für PG-Routen gibt es hingegen eine leichte Verbesserung gegenüber den PK-Routen. Dies ist allerdings für die Dimension der Zeit auch anzunehmen, da die schnellste Route auf Basis der personalisierten Kantengewichte immer eine bessere oder gleich gute Lösung berechnen sollte. Dagegen ist die schnellste Route auf Basis

der Standardkantengewichte meist schlechter als die kombinierte Personalisierung. Besonders ausgeprägt ist die Verschlechterung erneut für die sporadischen Fahrer mit einem Median von -1.5%, so dass anzunehmen ist, dass die personalisierten Kantengewichte für diese Klasse eine wichtige Rolle spielen.

Im nächsten Schritt wird die analoge Auswertung für die Dimension des Raumes durchgeführt, siehe Abbildung 7.29. Hier kommt es für alle Gruppen und alle Varianten für nahezu alle Fahrer zu einer Verschlechterung. Für PG- und PZ-Routen liegt die Verschlechterung im Median bei -1%, allerdings kann sie auch deutlich größer, insbesondere für die PG-Routen und für sporadische Fahrer, ausfallen (Median -2%).

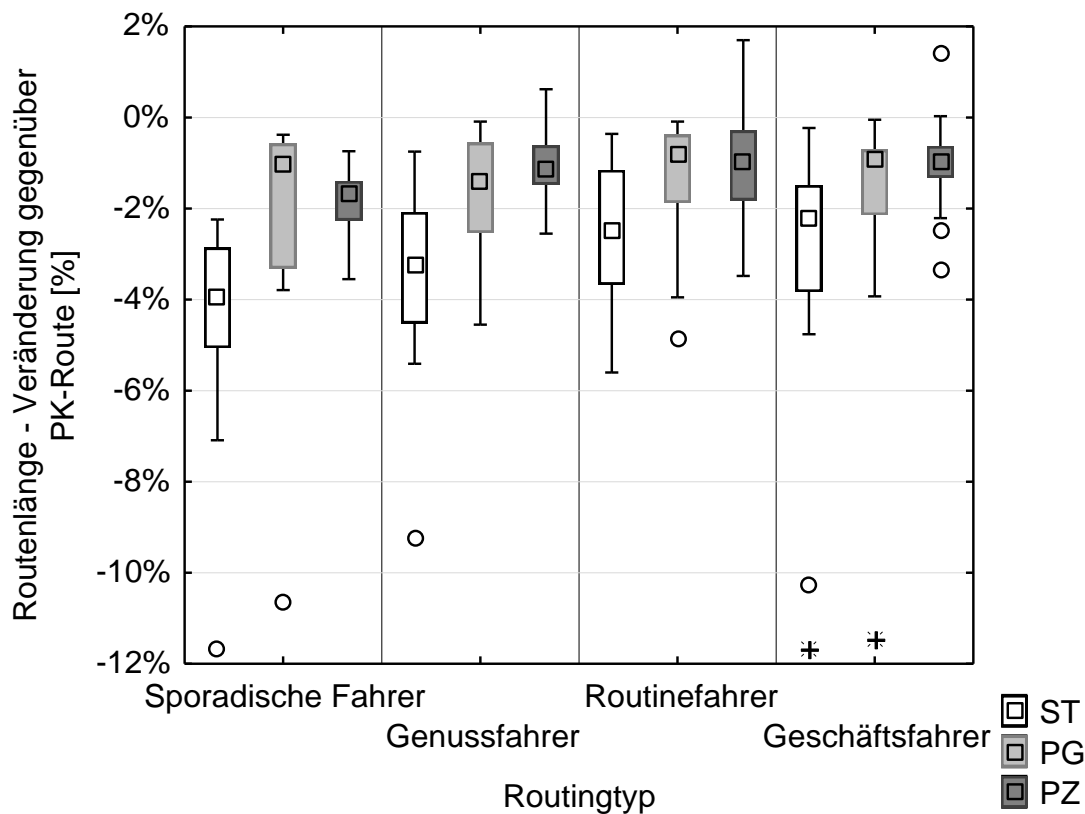


Abbildung 7.29: Bewertung der Routenlängenabweichung

Die Verschlechterung bei der ST-Route ist generell größer. Für Geschäftsfahrer liegt der Median bei -2%. Für Routine- und Genussfahrer befindet sich der Median knapp unter bzw. über -2%. Die größte Verschlechterung kommt erneut bei den sporadischen Fahrern vor. Hier liegt der Median bei -4%. Kein Fahrer hat eine geringere Verschlechterung als -2% und ein Viertel der Fahrer hat eine Verschlechterung von mehr als -5%.

Die Ergebnisse zeigen, dass der Einfluss der Personalisierung auf die Dimension des Raumes deutlich größer ist. Dieses war generell zu erwarten, da durch das Vorgehen in dieser Arbeit die Dimension des Raumes mit einbezogen und modelliert wird. Entscheidend ist im weiteren Schritt die Aggregation beider Dimensionen, um ein Gesamtbild der Verbesserung durch die Personalisierung zu erhalten.

Die Aggregation erfolgt, wie in Abschnitt 5.5.1 beschrieben, mit Hilfe der ermittelten Maße M_z und M_r für jeden Fahrer. Auf diese Weise wird ein Wert, der beide Dimensionen mit einbezieht, generiert und kann analysiert werden. Die Ergebnisse sind in Abbildung 7.30 dargestellt. Es zeigt sich erneut, dass ein Informationsverlust zu einer Verschlechterung für alle Klassen führt.

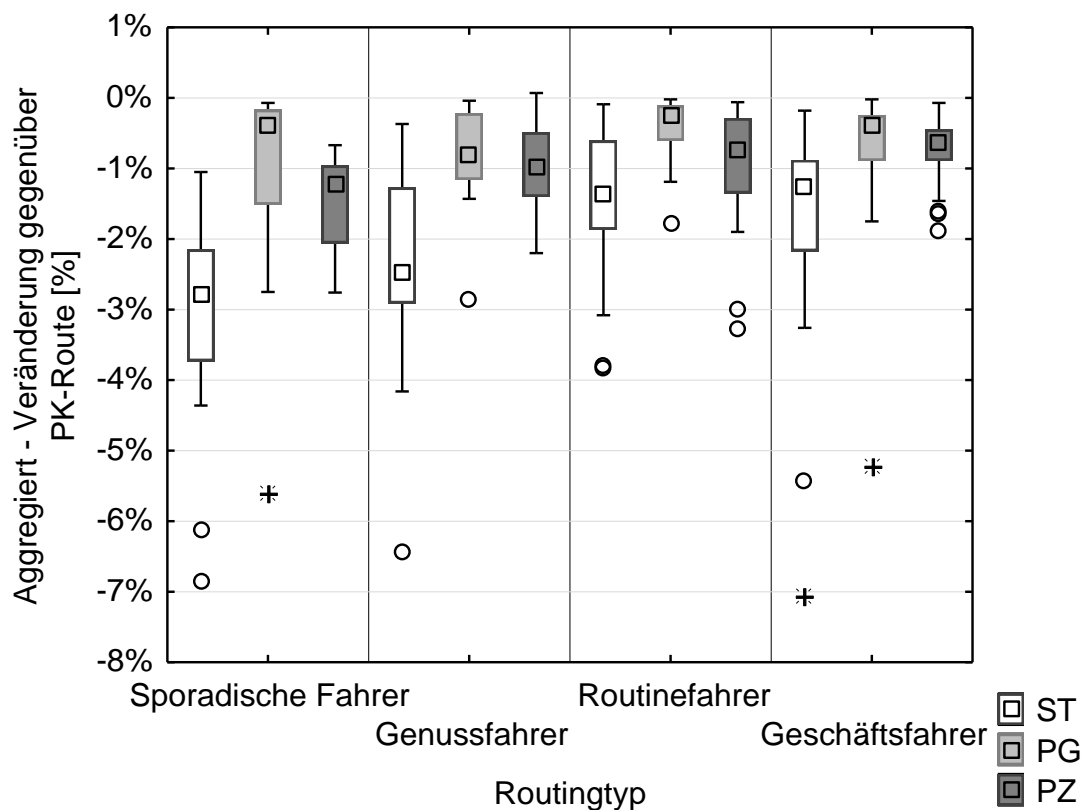


Abbildung 7.30: Bewertung der aggregierten Zielkriterienabweichung

Für PZ-Routen zeigen sich die bereits bekannten Muster. Für nahezu alle Fahrer gibt es eine Verschlechterung, die für sporadische Fahrer am stärksten ausgeprägt ist. Dabei liegt der Median zwischen -0.5 und -1%. In einem ähnlichen Bereich liegen die Mediane für PG-Routen. Hier ist allerdings die größte Verschlechterung für Genussfahrer vorhanden. Die sporadischen Fahrer haben nur eine geringe Verschlechterung zu verzeichnen. Dies liegt vor allem an der Dimension der Zeit, die

hier sogar zu einer Verbesserung führte und den negativen Effekt für die Dimension des Raumes nahezu ausgleicht.

Die größte Veränderung zeigt sich erneut bei der ST-Route. Hier liegt der Median für die Geschäfts- sowie die Routinefahrer unter -1% und für Genuss- und sporadische Fahrer zwischen -2% und -3%. Dazu gibt es Fahrer, bei denen der Informationsverlust zu einer deutlich größeren Verschlechterung von bis zu -7% führt. Dies zeigt das Verbesserungspotenzial insgesamt für die Personalisierung auf.

Abschließend wird noch der Einfluss der Dimension des Raumes und der entsprechenden Maße näher untersucht. Dazu wird analog zum vorherigen Vorgehen Maße und damit entsprechend Informationen für die Bestimmung der Zielfunktion weggelassen. In ersten Fall wird das Maß M_a für die Bestimmung der Gewichte für die Dimension des Raumes nicht genutzt (PK(1&2a,2b)). Im zweiten Fall werden zusätzlich die Ergebnisse der Kantenperspektive und das entsprechende Maß M_v weggelassen (PK(2a,2b)). Damit basieren die Ergebnisse ausschließlich auf den Analysen der Routenperspektive. Im dritten Fall wird nur die Netzwerkperspektive für die Dimension des Raumes genutzt (PK(1&2a,3)). Die Ergebnisse dieser Untersuchung sind in Abbildung ohne die Extremwerte 7.31 aufgetragen.

Die ausschließliche Nutzung der Netzwerkperspektive (PK(1&2a,2b)) führt zu unwesentlichen Veränderungen. Die Güte ist sehr ähnlich, wobei die Streuungen innerhalb der Klassen leicht unterschiedlich sind. Dagegen gibt es in den Fällen, in denen die Ergebnisse der Netzwerkperspektive nicht berücksichtigt wurden, größere Abweichungen in Abhängigkeit der Klassen. Für Geschäftsfahrer gibt es eine leichte Verschlechterung. Bei Routinefahrern hingegen gibt es nahezu keine Veränderungen. Im Gegensatz dazu gibt es bei Genussfahrern eine deutlichere Verschlechterung. Der Median für sporadische Fahrer ist zwar nur leicht negativ, doch zeigt die Whisker-Box an, dass für die meisten Fahrer das Weglassen der Informationen der Netzwerkperspektive zu einer Verschlechterung führt.

Generell sind die Auswirkungen bei weitem nicht so stark ausgeprägt wie bei den vorherigen Analysen, doch zeigen die Auswertungen, dass die Analysen der Netzwerkperspektive einen positiven Effekt auf das Ergebnis haben. Damit stellen die entsprechenden Analysen für die Dimension des Raumes einen wichtigen Input für die Modellierung und die daran anschließende Personalisierung dar.

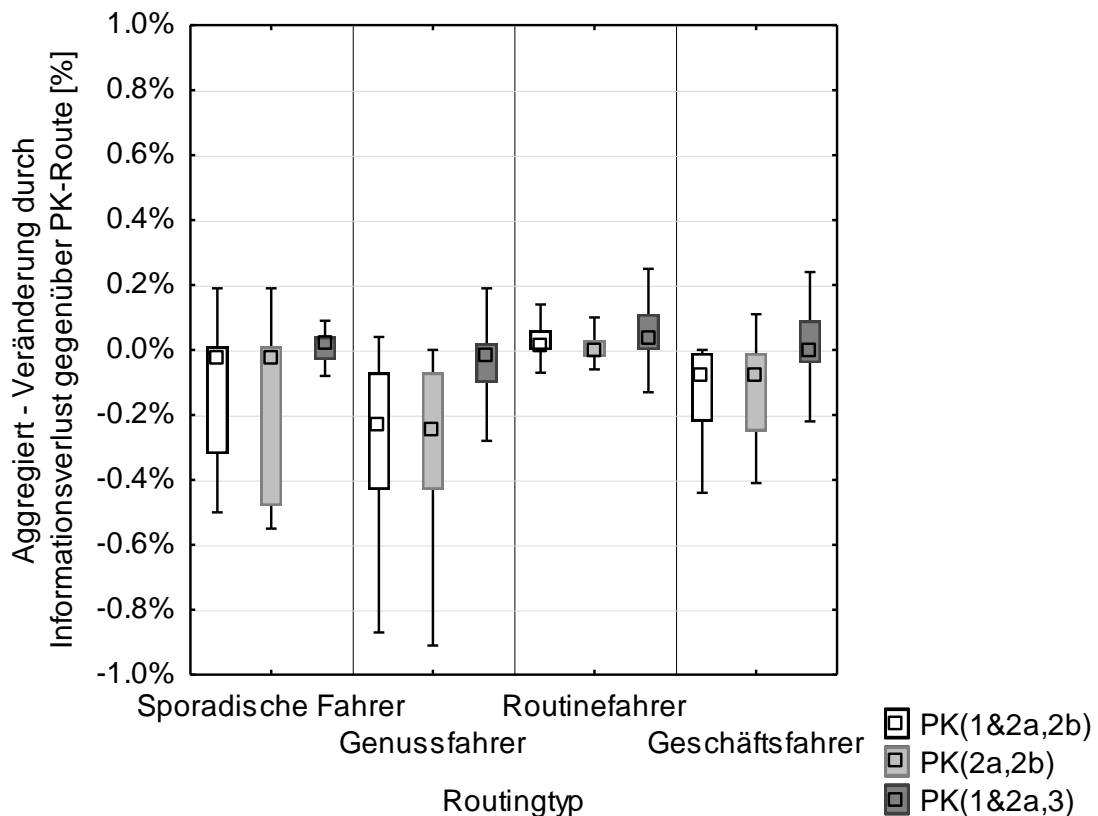


Abbildung 7.31: Einfluss der Perspektivenmodelle

7.6.2 Nutzung alternativer Gütekriterien

In diesem Abschnitt werden zwei weitere Gütekriterien betrachtet. In einem ersten Schritt werden die Anteile der Straßenklassen analysiert. Eine Personalisierung sollte dazu führen, dass die Anteile und damit die Streckencharakteristika der berechneten Routen in größerem Maße denen der gemessenen Routen entsprechen. So sollte einem Fahrer, der wenig auf der Autobahn fährt, auch eine Route mit geringeren Autobahnanteilen angeboten werden.

In einem zweiten Schritt werden die Routenverläufe direkt miteinander verglichen und die Anteile bestimmt, auf denen sie übereinstimmen. Eine Personalisierung sollte zur Folge haben, dass die berechneten Routen in einem höheren Maße mit den gemessenen übereinstimmen.

Analyse der Streckencharakteristika

Die Anteile der Straßenklassen werden kantenanzahl- sowie längenabhängig untersucht. Wie in Abschnitt 5.5.2 beschrieben, werden die Anteile der drei Straßenklas-

sen für alle Personalisierungsansätze berechnet und in Beziehung zur gemessenen Route gesetzt. In Abbildung 7.32 sind die Abweichungen für die verschiedenen Ansätze sowie Fahrerklassen dargestellt. Im Gegensatz zum vorherigen Maß der Routenkriterien gibt es hier keine eindeutige Unterscheidung für die Klassen. Bei hoher Streuung gibt es nur die Tendenz, dass Fahrer mit niedriger Präferenz für die Dimension des Raumes (sporadische Fahrer und Routinefahrer) auch geringere Abweichungen haben als Fahrer mit hoher Präferenz für die Dimension des Raumes (Genuss- und Geschäftsfahrer). Die Mediane liegen für alle Klassen und für alle Gruppen zwischen 4 und 6%.

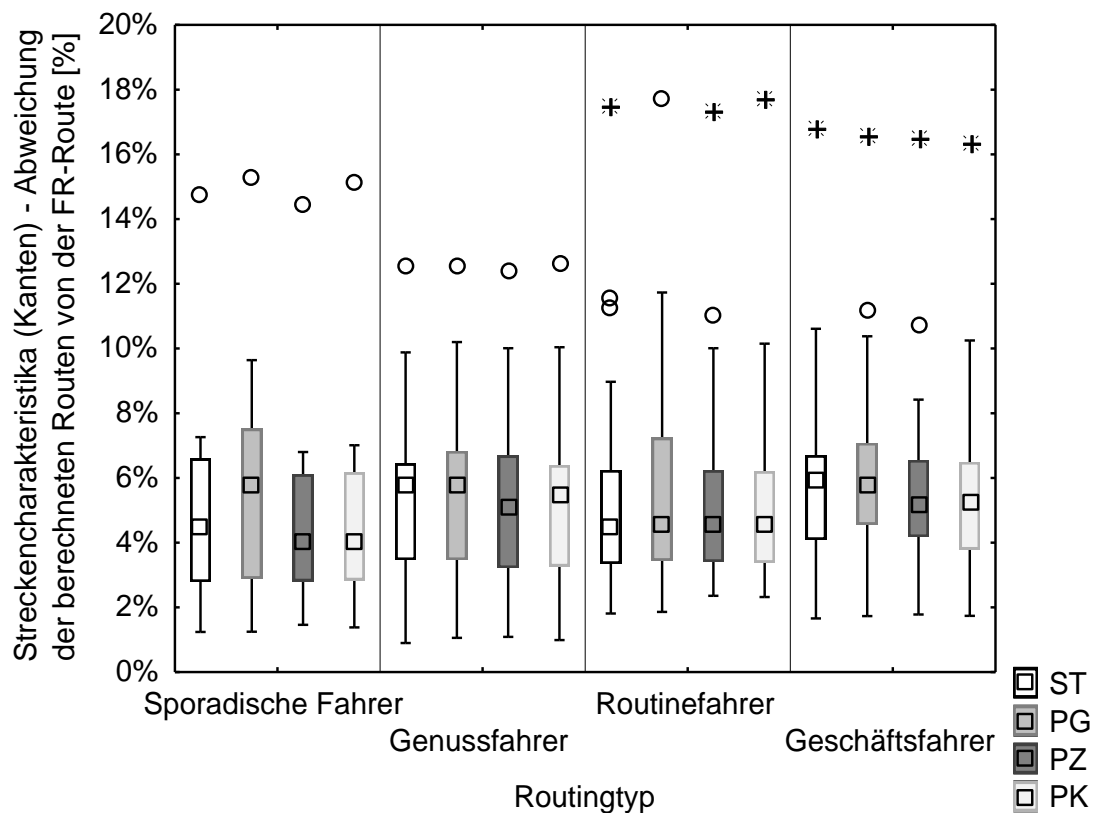


Abbildung 7.32: Bewertung der berechneten Routen für die kantenanzahlbasierten Straßenklassenanteile

Ein sehr ähnliches Bild zeigt sich, wenn man die Längen der Kanten berücksichtigt, siehe Abbildung 7.33. Der Median liegt bei 5% und Unterschiede zwischen den einzelnen Gruppen sind nicht feststellbar. Gleiches gilt auch für die Unterschiede zwischen den verschiedenen Ansätzen, was im Folgenden analog zum vorherigen Vorgehen untersucht wird.

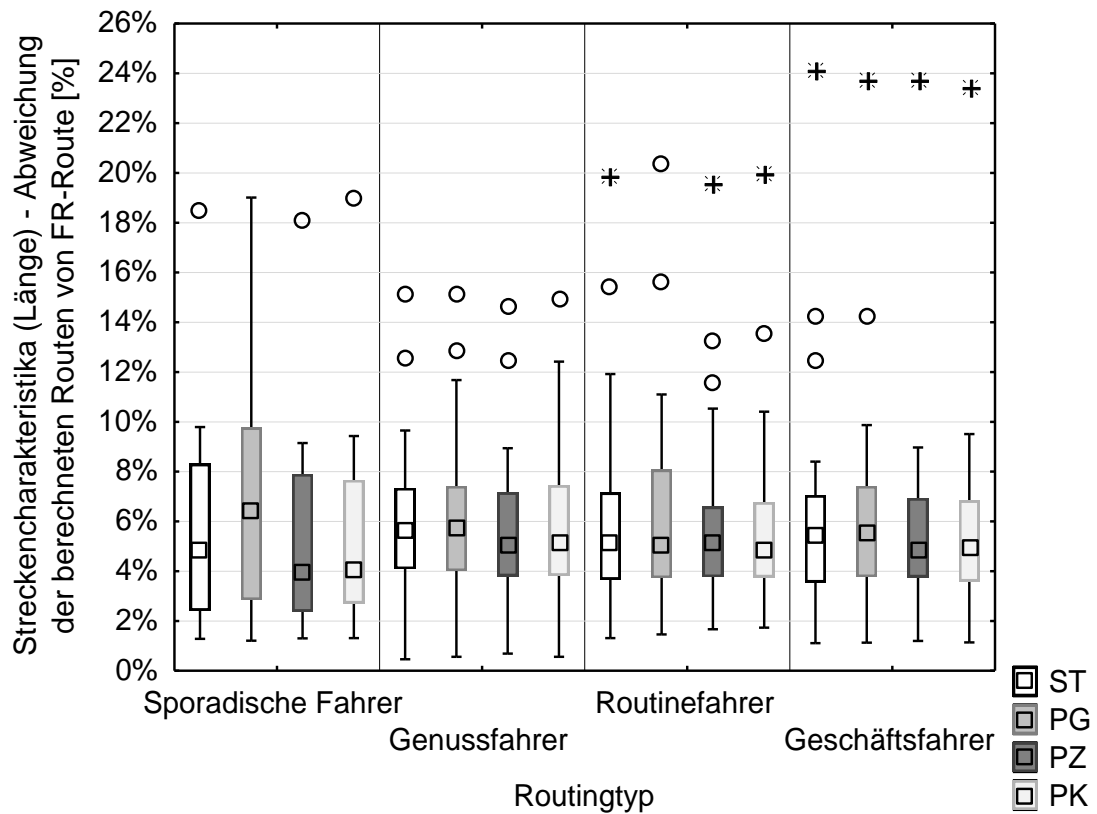


Abbildung 7.33: Bewertung der berechneten Routen für die längenbasierten Straßenklassenanteile

Um den Nutzen der Personalisierung darzustellen, wird die relative Veränderung in Bezug auf die PK-Route untersucht, siehe Abbildung 7.34. Für alle Klassen verändert sich die Abweichung für PZ-Routen nur unwesentlich. Der Median liegt nahe Null. Für PG-Routen hingegen ist der Median immer negativ, d.h., dass es zu einer größeren Abweichung kommt als dies bei der kombinierten Personalisierung der Fall ist. Insbesondere für sporadische Fahrer ist dieser Effekt sehr ausgeprägt. Für die ST-Route kann ebenfalls eine negative Tendenz beobachtet werden, wobei sich die Klassen etwas unterscheiden. Hier ist der größte negative Effekt bei den Genussfahrern zu finden. Dazu liegt der Median bei Routinefahrern wie für PZ-Routen nahe Null.

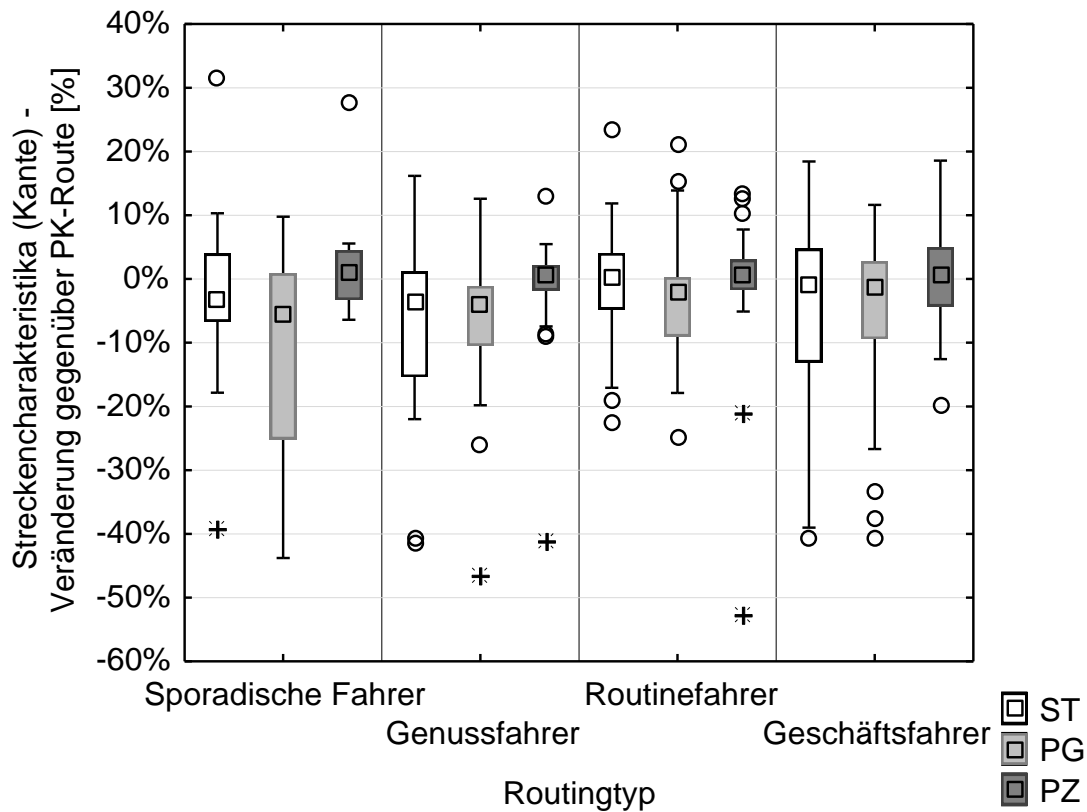


Abbildung 7.34: Nutzen der Personalisierung für die kantenbasierten Straßenklassenanteile

Ebenfalls ist zu erwähnen, dass die Abweichungen, die als Basis für den relativen Vergleich dienen, absolut sehr klein sind und somit zu sehr hohen Werten führen können. Dies zeigt sich insbesondere bei den Extremwerten. Allerdings ist auch hier das Ergebnis, dass ein Informationsverlust, wie schon bei den Zielkriterien, zu einer Verschlechterung führt. Dies hängt von der Einbeziehung der Dimension des Raumes ab. Wird diese auf den standardisierten Kantengewichten einbezogen (PZ-Route), so gibt es keine Verschlechterung.

Ein sehr ähnliches Bild zeigt sich, wenn die Längen der Kanten berücksichtigt werden. Die Ergebnisse sind in Abbildung 7.35 aufgetragen. Die Nichtberücksichtigung der Dimension des Raumes führt auch in diesem Fall zu einer Verschlechterung. Somit führen die Personalisierung und insbesondere die kombinierte Personalisierung zu einer Angleichung der berechneten an die gemessene Route bezüglich der Streckencharakteristika.

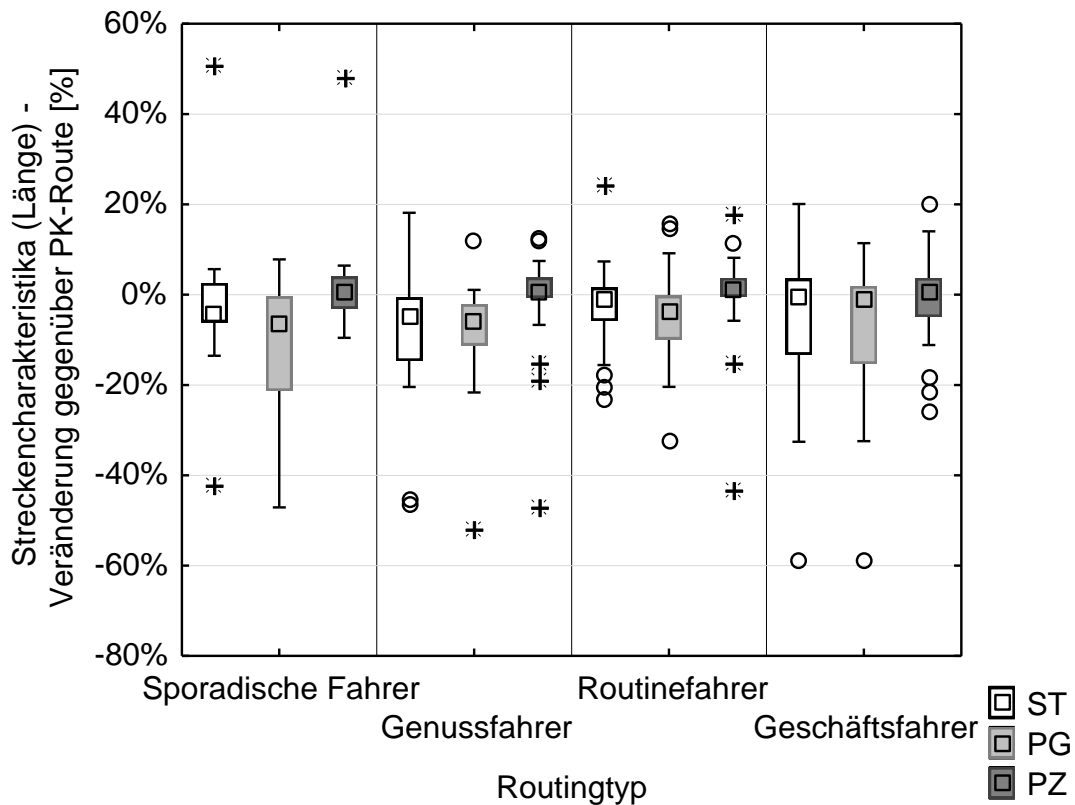


Abbildung 7.35: Nutzen der Personalisierung für die längenbasierten Straßenklassenanteile

Bestimmung der Übereinstimmung zwischen der gemessenen und den verschiedenen berechneten Routen

Abschließend wird die Übereinstimmung der berechneten Routen mit der gemessenen Route analysiert. Aufgrund der analogen Ergebnisse auf Kanten- sowie auf Längenbasis für die Straßenklassenanteile erfolgt dies ausschließlich auf Basis der Kanten. Es wird analysiert, wie viele Kanten des gemessenen Routenverlaufs auch für den berechneten Routenverlauf vorkommen und dann die Übereinstimmung ermittelt.

Die Übereinstimmung als Prozentanteil auf Basis der gemessenen Route ist in Abbildung 7.36 unterschieden nach Klassen dargestellt. Im Gegensatz zu den Anteilen der Straßenklassen wird hier erneut ein Klasseneffekt deutlich. Insbesondere die sporadischen Fahrer grenzen sich mit einer geringen Übereinstimmung ab. Hier liegt der Median bei 40%.

Dagegen liegt der Median für Genuss- und Routinefahrer bei knapp 60%, wobei eine große Streuung vorhanden ist. Für Geschäftsfahrer liegt der Median bei knapp

70% mit etwas geringerer Streuung. Damit bildet sich die Modellierung auch für dieses Gütemaß ab. Ob die Personalisierung zu einer Verbesserung führt, lässt sich aus der Darstellung hingegen nicht ableiten.

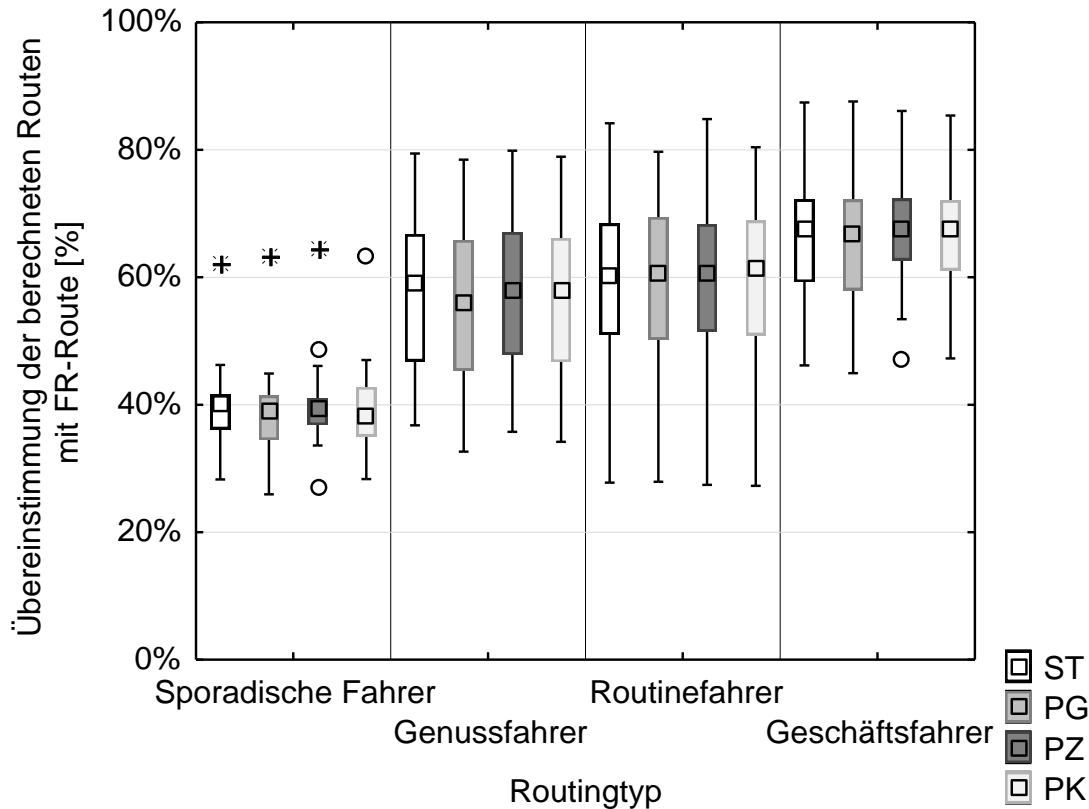


Abbildung 7.36: Bewertung der berechneten Routen für die Übereinstimmung

Dafür wird erneut die Veränderung im Vergleich zur kombinierten Personalisierung untersucht, siehe Abbildung 7.37. Für die PZ-Routen liegt der Median für Routine- und Geschäftsfahrer nahe Null, für sporadische und Genussfahrer hingegen liegt er zwischen 1 und 3%. Das bedeutet, dass der Ausschluss der personalisierten Kantengewichte zu einer Verbesserung führt.

Für die PG-Routen liegt der Median für alle Klassen nahe Null. Es ist nur für sporadische Fahrer eine leichte Tendenz erkennbar, dass die Ergebnisse sich für die Klasse verschlechtert. Somit kommt es durch den Informationsverlust zu keiner Verschlechterung. Dies zeigt sich auch für die ST-Route. Hier kommt es sogar erneut für sporadische und Genussfahrer zu einer Verbesserung (Median liegt zwischen 1 und 2%), obwohl weder angepasste Kantengewichte noch eine personalisierte Zielfunktion genutzt werden.

Somit lässt sich für das Gütemaß der Übereinstimmung festhalten, dass für sporadische und Genussfahrer die Hinzunahme der personalisierten Zielfunktion (PG- und PK-Routen) zu einer Verschlechterung führt. Für Routine- und Geschäftsfahrer verändert sich die Übereinstimmung durch die Personalisierung nicht.

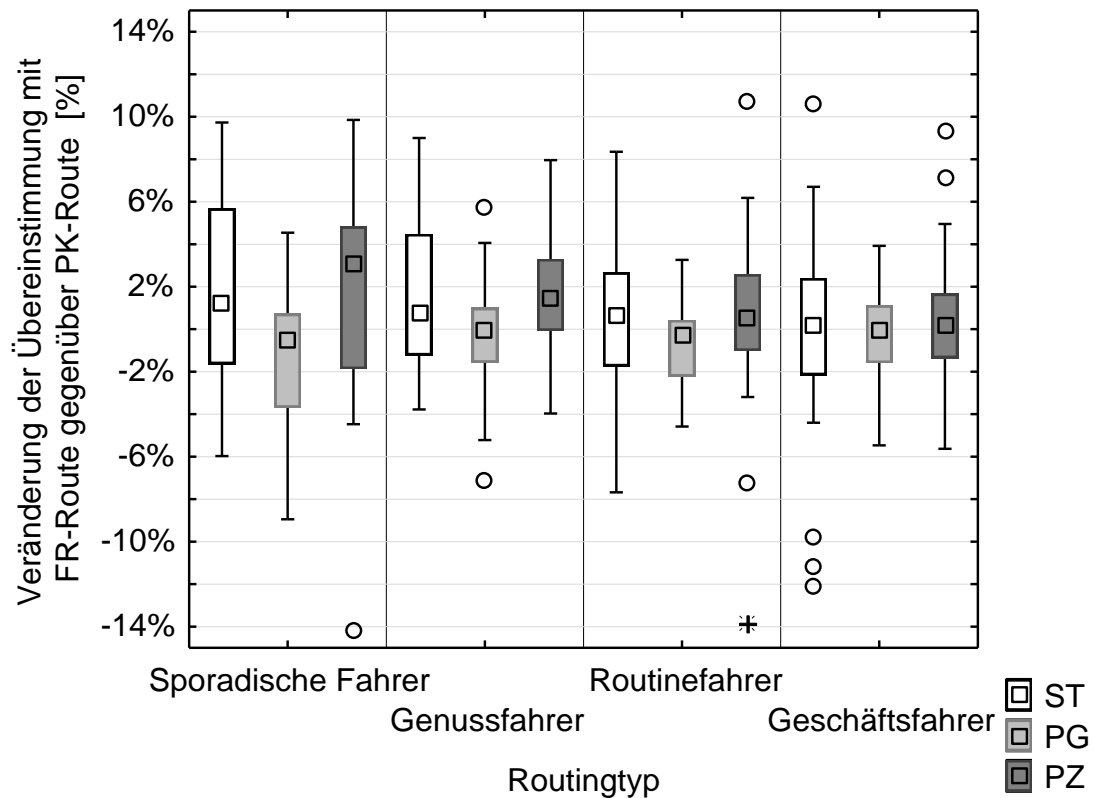


Abbildung 7.37: Nutzen der Personalisierung für die Übereinstimmung

7.6.3 Diskussion der Bewertungsergebnisse

Um den Nutzen der Personalisierung darzustellen, wurden die berechneten Routen mit Hilfe verschiedener Gütemaße bewertet. Ziel war es, Routen zu berechnen, die in größerem Maße die navigatorischen Präferenzen der Fahrer berücksichtigen. Dies wird durch die verschiedenen Gütemaße abgebildet. Die Nutzung der Zielkriterien bildet dabei das direkte Maß zur Bewertung einer Route und ist auch Gegenstand der vorherigen Analysen und Modellierung.

Für dieses Gütemaß konnte eine Verbesserung durch die Personalisierung gezeigt werden. Dafür wurde dargestellt, dass ein Informationsverlust durch eine weniger umfassende bzw. keine Personalisierung zu einer Verschlechterung führt. Es konnte auch gezeigt werden, dass der Verlust einer Information (Personalisierung der

Kantengewichte oder der Zielfunktion) zu einer Verschlechterung führt und somit alle in dieser Arbeit generierten Informationen relevant sind.

Auch mit Hilfe des zweiten Maßes der Streckencharakteristika konnte der Nutzen der Personalisierung aufgezeigt werden. Eine Personalisierung führt zu einer Route, für die die Abweichung der Straßenklassenanteile mit der gemessenen Route geringer ist als für die anderen berechneten Routen. Die Streckencharakteristika sind kein direktes Ziel der Routenberechnung, somit werden diese nur indirekt durch die Personalisierung beeinflusst. Für diese Maße kann somit eine Verbesserung der Routingkomponente durch die Personalisierung quantifiziert werden.

Es wäre zu erwarten gewesen, dass durch die Berücksichtigung der Präferenzen Routen berechnet werden, die mit denen von Fahrern ausgewählten übereinstimmen. Hierfür wurde das Gütemaß der Übereinstimmung genutzt und die berechneten Routen dahingehend analysiert. Ein Nutzen der Personalisierung konnte mit Hilfe dieses Maßes allerdings nicht gefunden werden.

Die Gründe hierfür sind vielschichtig. So ist der gesamte Ansatz der Personalisierung global aufgebaut und behandelt alle Kanten als Klassen. Damit werden auch vom Fahrer häufig befahrene Kanten nicht bevorzugt behandelt und eine ähnliche Parallelstrecke kann ausgewählt werden. Dies kann auch durch die für die Personalisierung genutzte Karte verursacht sein, in der Unterschiede auf dieser Ebene (z.B. Manöverstrafen), die zu einer Bevorzugung bestimmter Kanten führen, nicht abgebildet sind. Dies zeigt sich auch in der generell niedrigen Übereinstimmung der berechneten Routen mit der FR-Route von 40 bis 70%.

Der Nutzen der Personalisierung besteht darin, dass die Zielkriterien in größerem Maße erreicht werden unter der Annahme, dass diese den Präferenzen des Fahrers entsprechen. Zwar stimmen vom Fahrer ausgewählte Routen und personalisierte Routen nicht überein, doch zeigt sich, dass sie ähnlicher werden und damit den Präferenzen des Fahrers in höherem Maße entsprechen. Es ist anzunehmen, dass die Personalisierung damit auch zu einer höheren Akzeptanz der berechneten Routen durch den Fahrer führt, was zu prüfen wäre.

8 Fazit und Ausblick

In dieser Arbeit wurde ein Verfahren für die Personalisierung der Routingkomponente von Navigationssystemen auf Basis eines Modells der navigatorischen Präferenzen eines Fahrers entwickelt und vorgestellt. Die Modellierung umfasste sowohl die Dimension der Zeit wie auch die Dimension des Raumes, welche als die relevanten Dimensionen bei der Routenauswahl ermittelt wurden. Aufbauend auf dem Modell wurden verschiedene Varianten der Personalisierung entwickelt und der Nutzen dargestellt und bewertet.

Für die Implementierung des Verfahrens wurde ein Realfahrdatensatz aus dem Projekt euroFOT genutzt. Er umfasste das Routenauswahlverhalten von ca. 100 Fahrern über einen Zeitraum von drei Monaten. Damit ermöglichte der Datensatz aufgrund seiner Größe aber auch der Qualität der Daten die Analyse des Routenauswahlverhaltens von Fahrern. Dafür wurden in dieser Arbeit mehrere Datenaufbereitungsschritte durchgeführt. Diese umfassten die Datenauswahl, die Datenbereinigung, insbesondere des Matchings, und die Datenreduktion.

Für die Erstellung eines zweidimensionalen Modells des Fahrers wurde das Routenauswahlverhalten für drei Perspektiven untersucht: Die Kantenperspektive, die Routenperspektive und die Netzwerkperspektive. Für die Kantenperspektive wurden die vom Fahrer gefahrenen Geschwindigkeiten gemessen und mit denen in der digitalen Karte gespeicherten verglichen. Der aggregierte Vergleich über mehrere Klassen führte zu Modell 1 des Fahrers, das der Dimension der Zeit zugeordnet wurde.

Ebenfalls wurden die gemessenen Geschwindigkeiten für die Personalisierung genutzt. Dieser Ansatz verlief analog zum Vorgehen in heutzutage verfügbaren Navigationssystemen. Dabei zeigte es sich, dass die in der digitalen Karte hinterlegten Geschwindigkeiten den durchschnittlichen Geschwindigkeiten der Fahrer entsprechen, es aber große Unterschiede zwischen den Fahrern, insbesondere bei höheren Geschwindigkeitsklassen, gibt.

Das Ergebnis der Analysen für die Routenperspektive waren zwei Modelle 2a und 2b. Hierfür wurden für die FR-Routen die Start- und Zielpositionen bestimmt.

Für die Positionen wurde mit Hilfe einer neutralen externen Routingreferenz eine schnellste Route berechnet. Diese berechnete Route diente als Benchmark, und es wurden die Routendauer und die Routenlänge der gemessenen Route des Fahrers mit der entsprechenden berechneten Route verglichen. Die erhaltenen Vergleichswerte für die Routen wurden für jeden Fahrer zusammengefasst. Dabei konnte der Wert für die Routendauer der Dimension der Zeit und der für die Routenlänge der Dimension des Raumes zugeordnet werden.

Für die Netzwerkperspektive wurde das vom Fahrer aufgespannte Netzwerk durch einen Graphen repräsentiert. Dieser Graph wurde mit Analysewerkzeugen bezüglich seiner Größe und seiner Struktur untersucht und Kennwerte für diese Indikatoren bestimmt. Das Ergebnis war die Zuordnung eines jeden Fahrers zu einem Navigationstyp, der als Modell 3 für die weiteren Schritte genutzt und der Dimension des Raumes zugeordnet wurde. Insbesondere dieses Modell 3 stellte sicher, dass die Dimension des Raumes in der Modellierung in ausreichendem Maße berücksichtigt wurde.

Das Ergebnis der Analysen für die verschiedenen Perspektiven waren jeweils zwei Modelle für die Dimension der Zeit und zwei für die Dimension des Raumes. Eine graphische sowie statistische Analyse zeigte, dass es nicht sehr stark ausgeprägte aber signifikante Korrelationen zwischen Modellen der gleichen Dimension und Modellen der gleichen Perspektive gab. Modelle unterschiedlicher Perspektive und Dimension hingegen zeigten keine Korrelationen. Somit konnten zwar die Modelle gleicher Dimension aggregiert werden, jede Dimension für sich aber als eigenständige Teil in die zweidimensionale Modellierung der navigatorischen Präferenzen eines Fahrers eingehen.

Auf Basis der Kennwerte für die Dimension der Zeit und der Dimension des Raumes wurde für die Modellierung eine Klassifizierung mit Hilfe eines Clusterverfahrens durchgeführt und so die Fahrer einem Routingtyp zugeordnet. Fahrer mit hohen Werten für beide Dimensionen wurden als Geschäftsfahrer und Fahrer mit kleinen Werten für beide Dimensionen als sporadische Fahrer klassifiziert. Routinefahrer hatten hohe Werte für die Dimension der Zeit und niedrige für die Dimension des Raumes. Im umgekehrten Fall wurden die Fahrer als Genussfahrer klassifiziert.

Für die anschließende Personalisierung wurden drei verschiedene Ansätze entwickelt. Der erste Ansatz umfasste die Personalisierung der Kantengewichte, die auf Basis der Analyseergebnisse der Kantenperspektive an die Fahrer angepasst wurden. Auf dieser Basis wurde eine schnellste Route mit Hilfe einer internen Routingreferenz, die eine digitale Karte auf Basis des Realfahrdatensatzes nutzte, berech-

net. Neben der Flexibilität, die dieses Vorgehen ermöglichte, war somit auch die Vergleichbarkeit mit den Routen der Fahrer gesichert.

Im zweiten Ansatz wurde die Zielfunktion personalisiert, indem aus dem Modell Anteile für die Dimension der Zeit und der Dimension des Raumes bestimmt wurden. Für diese zweidimensionale Zielfunktion wurde auf Basis der Standardkantengewichte eine personalisierte Route berechnet. Der dritte Ansatz kombinierte die beiden vorherigen Ansätze und berechnete eine Route für eine zweidimensionale Zielfunktion auf Basis der angepassten Geschwindigkeiten.

Um den Nutzen der verschiedenen Ansätze der Personalisierung aufzuzeigen, wurde eine Standardroute berechnet. Dies war eine schnellste Route auf Basis der Standardkantengewichte. Für den Vergleich der Standardroute mit den personalisierten Routen wurden verschiedene Gütemaße genutzt. Das erste Gütemaß untersuchte die Zielkriterien der Routen. Die Routendauern und Routenlängen wurden zuerst getrennt und danach aggregiert analysiert. Die Werte für das Gütemaß der Zielkriterien waren für die Routen der kombinierten Personalisierung am Höchsten. Im Vergleich zu diesem Ansatz führte der Verlust von Informationen zu einer Verringerung der Zielkriterienerreichung, die insbesondere für die Standardroute aber auch für sporadische sowie Genussfahrer stark ausgeprägt war.

Ebenfalls konnte gezeigt werden, dass die Einbeziehung des Navigationstyps (Modell 3) und damit die Dimension des Raumes ein wichtiger Faktor für die Güte darstellt. Damit war die Modellierung der Dimension des Raumes ein wichtiger Schritt der Personalisierung. Beide Dimensionen, sowohl Zeit wie auch Raum, wurden bei der Routenauswahl durch einen Fahrer einbezogen, was in dieser Arbeit auch bei der Routingkomponente von Navigationssystemen der Fall war.

Als eines von zwei Alternativgütemaßen wurden die Streckencharakteristika genutzt. Es wurden die Anteile für die Anteile der Straßenklassen Innerorts-, Überland- und Autobahnstrecken sowohl für die gemessene Route der Fahrer wie auch für die berechneten Routen bestimmt. Dabei waren die Anteile der personalisierten berechneten Routen vergleichbarer mit der gemessenen Route als die Anteile der Standardroute mit der gemessenen Route. Besonders ausgeprägt war der Effekt erneut für die kombinierte Personalisierung sowie bei sporadischen Fahrern. Damit wurden die berechneten Routen durch die Personalisierung ähnlicher zu denen vom Fahrer ausgewählten, und es ist anzunehmen, dass sie damit die navigatorischen Präferenzen in höherem Maße berücksichtigen und zu einer höheren Akzeptanz führen.

Ein ähnliches Vorgehen wurde beim dritten Gütemaß der Übereinstimmung verfolgt. Hierbei wurden nicht nur die Charakteristika sondern der Routenverlauf an sich miteinander verglichen. Es wäre zu erwarten gewesen, dass eine personalisierte Route in höherem Maße mit der vom Fahrer ausgewählten Route übereinstimmt, was sich allerdings nicht zeigen ließ. Zwar wurde die Route durch die Personalisierung verändert, was aber nicht zur Folge hatte, dass vermehrt Kanten genutzt werden, die auch der Fahrer genutzt hat. Dies war allerdings zu erwarten, da die Personalisierung auf einer generellen klassenbasierten Ebene durchgeführt wurde und sich damit Veränderungen nur auf Klassenebene auswirkten, was für das zweite Maß auch gezeigt wurde.

Somit konnte für zwei Gütemaße eine Verbesserung der Routingkomponente durch eine Personalisierung quantifiziert werden. Das in dieser Arbeit entwickelte Verfahren stellt somit eine Möglichkeit dar, fundierte Informationen über das Routenauswahlverhalten des Fahrers zu generieren und in einem Modell der Routingkomponente für eine Personalisierung zur Verfügung zu stellen. Damit konnte auch die Machbarkeit eines solchen Vorgehens gezeigt werden.

Das Verfahren nutzte für die Analysen den gesamten Zeitraum von drei Monaten der euroFOT-Projektlaufzeit. Ein im Fahrzeug anwendbares Verfahren kann für die Kanten- und Routenperspektive direkt darauf aufbauen. Welcher Zeitraum notwendig ist, um das Modell der Netzwerkperspektive zu erstellen, wäre zu evaluieren. Die Ergebnisse zeigten, dass die Einbeziehung der Netzwerkperspektive zu einer besseren sowie robusten Lösung führt.

Die generierten Informationen über den Fahrer können auch weiteren Systemen im Fahrzeug zur Verfügung gestellt werden [50]. Für Systeme, die sich an das Fahrverhalten anpassen, z.B. adaptive FAS, ist von großer Bedeutung, welche Faktoren das Fahrerverhalten beeinflussen. Erste Untersuchungen auf Basis der euroFOT-Daten haben gezeigt, dass sich das Verhalten der Fahrer auf bekannten Strecken von dem auf unbekannten Strecken unterscheidet [47]. Beim Geschwindigkeitsverhalten sind diese Effekte nicht so stark ausgeprägt und wurden daher in dieser Arbeit ausgeklammert. Allerdings scheint der Abstand zum Vorderfahrzeug auf unbekannten Strecken größer zu sein als auf bekannten Strecken, was für das Beispiel des adaptiven Bremsassistenten aus Abschnitt 2.4.2 eine relevante Information darstellt.

An die Ergebnisse schließt sich ebenfalls die Frage nach weitergehenden Forschungsmöglichkeiten an. Die Modellierung beschreibt die generellen navigatorischen Präferenzen eines Fahrers. Aufbauend darauf kann das Modell zeitlich sowie räumlich differenziert werden. So kann es Zeiträume bzw. Fahrten geben, für die der Fah-

rer andere Präferenzen hat. Auf dem Weg zur Arbeit kann der Fahrer andere Präferenzen haben als auf einer Freizeitfahrt. Dies kann das Ergebnis der Modellierung sowie die Bewertung der personalisierten Route beeinflussen. Für die Ebene der Fahrzeugführung wird in der Forschung analog zwischen Fahrtyp und Fahrstil unterschieden und die Frage nach dem Zusammenhang ist eine zu klärende Frage.

Eine Möglichkeit der räumlichen Differenzierung ist die Berücksichtigung der Bekanntheit von Kanten. Wie bereits oben dargestellt, unterscheidet sich das Fahrverhalten auf bekannten von unbekannten Kanten. Daher könnten bekannte Kanten mit einem Bonus versehen werden, unter der Annahme, dass es für den Fahrer angenehmer ist auf bekannten Kanten zu fahren und er die Zielkriterien besser erreicht. Diese Annahme wäre zunächst zu evaluieren.

Abschließend wird noch auf den Verkehr als einen sehr wichtigen Einflussfaktor näher eingegangen. In Abschnitt 2.3.2 wurden bereits die kurzfristigen und langfristigen Einflüsse detailliert beschrieben. Diese beeinflussen ebenfalls das Routenauswahlverhalten des Fahrers (z.B. Umfahrung eines Staus), aber auch das Geschwindigkeitsverhalten (z.B. stehen im Stau). Somit hat der Verkehr sowohl Einfluss auf die Dimension der Zeit wie auch auf die Dimension des Raumes und sollte auch bei der Modellierung und der Personalisierung berücksichtigt werden. Dabei ist die Frage zu klären, ob und in welchem Ausmaß sich der Verkehr auf den Fahrer auswirkt.

Hierfür kann auf Arbeiten der Verkehrsforschung sowie auf den Erkenntnissen in dieser Arbeit aufgebaut werden. Ein kombinierter Ansatz zur Verbesserung der Navigation sollte sowohl Verkehr wie auch den Fahrer mit einbeziehen, um ein realistisches Modell des Fahrers und seiner navigatorischen Präferenzen zu erhalten, aber auch die Auswirkungen des Verkehrs richtig zu modellieren. Dies stellt einen weiteren entscheidenden Schritt für zukünftige Arbeiten zur Verbesserung der Navigation dar, für den in dieser Arbeit mit der Modellierung des Fahrers und seiner navigatorischen Präferenzen ein wichtiger Grundstein gelegt wurde.

Literaturverzeichnis

- [1] Ahrholdt, D. (2010). Erfolgsfaktoren einer E-Commerce Website: Empirische Identifikation vertrauensfördernder Signale im Internet-Einzelhandel, Gabler, Wiesbaden.
- [2] Anand, S. and Buchner, A. (1998). Decision Support Using Data Mining. Financial Times Pitman Publishers, London.
- [3] Angermann, A. et al (2011). MATLAB - Simulink - Stateflow, 7. Auflage, Oldenbourg Wissenschaft, München.
- [4] AreaMobile Redaktion (2007). Garmin löst TomTom als Marktführer ab, <http://www.areamobile.de/news/7796-navisysteme-garmin-loest-tomtom-als-marktfuehrer-ab>, Abgerufen am 09.01.16.
- [5] Arnett, J. (1994). Sensation seeking: A new conceptualization and a new scale. In: Personality and Individual Difference, 16(2), pp. 289-296. Elsevier Science Ltd., London.
- [6] Barnard, Y. and Carsten, O. (2010). Field Operational Tests: Challenges and methods. In: Krems J., Petzold T. and Henning M. (Ed) Proceedings of the European Conference on Human Centred Design for Intelligent Transport Systems, pp. 323-332, April 29-30, 2010, Berlin.
- [7] Bernard, M. and Axhausen, K. W. (2010). Ein neuer Ansatz für standardisierte Ganglinien. In: Arbeitsbericht Verkehrs- und Raumplanung 540, Juli 2010, IVT, ETH, Zürich.
- [8] Bernotat, R. (1970). Anthropotechnik in der Fahrzeugführung. In: Ergonomics, 13(3), pp. 353-377.
- [9] Bertele, C. (2014). Die Zukunft der Navigationsbranche. In: Telecom Handel, <http://www.telecom-handel.de/consumer-communications/navigation/die-zukunft-der-navigationsbranche-805598.html?ganzseitig=1>, Abgerufen am 18.07.15.

- [10] Bloch, J. (2006). How to Design a Good API and Why it Matters, <http://static.scribd.com/docs/908bil5xonqxe.pdf>, Abgerufen am 09.01.16.
- [11] Bonsall, P. W. and Parry, T. (1990). Drivers' requirements for route guidance. In: Proceedings of third international conference on Road Traffic Control, IEE, pp. 1-5. 1-3 May 1990, London.
- [12] Bortz, J. and Schuster, C. (2010). Statistik für Human- und Sozialwissenschaftler. Springer, Berlin Heidelberg.
- [13] Bouslimi, W. (2005). Augmented Naive Bayesian Network for Driver Behavior Modeling. In: Proceedings of Intelligent Vehicles Symposium, pp. 236-242, 6-8 Juni 2005 , Las Vegas.
- [14] Bubbs, H. (1993). Reliability of the Driver - A Method of Driver Modelling for Prevention of Driver Failure. In: Proceedings of 6th ICTCT Workshop, pp. 33-39, Oktober 1993, Salzburg.
- [15] Cabena, P. et al. (1998). Discovering Data Mining: From Concepts to Implementation, Prentice Hall Saddle River, New Jersey.
- [16] Cacciabue, P. C. (2007). Modelling Driver Behaviour in Automotive Environments, Springer, London.
- [17] Chia, A. (2013). Frost & Sullivan: Car Navigation Systems Market in 3 ASEAN Countries To Grow at a CAGR of 17.2 Per Cent To Reach 2.08 Million in 2018, <http://www.frost.com/prod/servlet/press-release.pag?docid=284811483>, Abgerufen am 09.01.16.
- [18] Cios, K. et al. (2000). Diagnosing myocardial perfusion from SPECT bullseye maps - a knowledge discovery approach. In: IEEE Engineering in Medicine and Biology Magazine, special issue on Medical Data Mining and Knowledge Discovery, 19(4), pp. 17-25.
- [19] Cole, D. J., Pick, A. J. and Odhams, A. M. C. (2006). Predictive and linear quadratic methods for potential application to modelling driver steering control. In: Vehicle System Dynamics: International Journal of Vehicle Mechanics and Mobility Volume 44, Issue 3, pp. 259-284.
- [20] Computerbild.de (2010). So kommen die Karten ins Navi, <http://www.computerbild.de/artikel/cb-Test-Navigation-Navi-Karten-Kartenmaterial-3002774.html>, Abgerufen am 09.01.16
- [21] Dia, H. and Panwai, S., (2007). Modelling Drivers compliance and route choice behaviour on response to travel information. In: Nonlinear Dynamics, Volume 49, Issue 4, pp. 493-509.

- [22] Dijkstra, E. W. (1959). A note on two problems in connexion with graphs. In: Numerische Mathematik, 1(1), pp. 269-271, Springer, Amsterdam.
- [23] Dirscherl, H.-C. (2014). Diese Verkehrslage-Dienste gibt es für Autofahrer. In: PC-Welt, <http://www.pcwelt.de/ratgeber/Stau-Warnung-Google-Maps-Tomtom-Verkehrslage-Echtzeitverkehrsinformationen-373385.html>, Abgerufen am 09.01.16.
- [24] Ehmke, J. F., Meisel, S. and Mattfeld, D. C. (2012). Floating Car Based Travel Times for City Logistics. In: Transportation Research - Part C: Emerging Technologies, 21(1), pp. 338-352, Elsevier Ltd.
- [25] Ehrenpfordt, I. and Rataj, J. (2006). Advanced Navigation Systems - Using Environment-based Workload to Individualize Route Planning. In: Proceedings 13th World Congress and Exhibition on Intelligent Transportation Systems and Services, ERTICO, London.
- [26] Ehrenpfordt, I. (2009). Zur belastungsoptimierten Routengenerierung im Straßenverkehr. In: Berichte aus dem DLR-Institut für Verkehrssystemtechnik, Volume 4, Braunschweig.
- [27] Fayyad, U., Piatetsky-Shapiro, G. and Smyth, P. (1996). From data mining to knowledge discovery: An overview. In: Advances in Knowledge Discovery and Data Mining, pp. 1-34, AAAI Press, Cambridge.
- [28] Felts, A. (2013). Das Ziel ist der Weg, <http://www.zeit.de/mobilitaet/2013-11/auto-navi-kartendaten>, Abgerufen am 09.01.16.
- [29] Finanzen.net (2014a). TomTom-Aktie, http://www.finanzen.net/bilanz_guv/TomTom, Abgerufen am 09.01.16.
- [30] Finanzen.net (2014b). Garmin-Aktie, http://www.finanzen.net/bilanz_guv/Garmin, Abgerufen am 09.01.16.
- [31] Finanznachrichten.de (2013). Telenav passt “Scout for Cars“ für den europäischen Automobilmarkt an - Personalisierte Navigations-, Verkehrs- und Suchfunktionen speziell für europäische Fahrer, <http://www.finanznachrichten.de/nachrichten-2013-11/28572224-telenav-passt-scout-for-cars-fuer-den-europaeischen-automobilmarkt-an-personalisierte-navigations-verkehrs-und-suchfunktionen-speziell-fuer-europ-256.htm>, Abgerufen am 09.01.16.

- [32] Friedrich, M. et al (2008). Mobile Phone Data for Telematic Applications. In: Proceedings of International Multi-Conference on Engineering and Technological Innovation: IMETI 2008, International Institute of Informatics and Systemics (IIS), Orlando, Florida.
- [33] Follmer, R. et al. (2008). Mobilität in Deutschland 2008 Ergebnisbericht Struktur - Aufkommen - Emissionen - Trends (No. Bundesministerium für Verkehr, Bau und Stadtentwicklung FE-Nr. 70.801/2006). Bonn & Berlin: infas & DLR.
- [34] Garmin Ltd. (2016). Live Traffic via DAB+, <https://www.garmin.com/de/pnd/dab-traffic>, Abgerufen am 09.01.2016.
- [35] Gediga, G. and Holling, H. (2015). Statistik - Testverfahren. Hogrefe, Göttingen.
- [36] Gesellschaft für Unterhaltungs- und Kommunikationselektronik (2010). Der Markt für Consumer Electronics - Deutschland 2010, [http://www.gfu.de/srv/easyedit/_ts_1316685518000/page:home/marktzahlen/markt/sl_1316619586955/args.link01/de_CE-Markt% 20-% 20Deutschland% 202010_mit_Text.pdf](http://www.gfu.de/srv/easyedit/_ts_1316685518000/page:home/marktzahlen/markt/sl_1316619586955/args.link01/de_CE-Markt%20-%20Deutschland%202010_mit_Text.pdf), Abgerufen am 09.06.15.
- [37] Google Inc. (2016). <https://www.google.de/maps/@52.2671125,10.508613,15z/data=!5m1!1e1?hl=de>, Abgerufen am 09.01.2016.
- [38] Gröschl, H. (2002). Galileo-Europas Antwort auf GPS, In Mensch & Natur [http://feeds.stgsrv.de/index.php?id=26&tx_ttnews\[tt_news\]=969&cHash=bb42b87b90ba01792e8d29e8e78aa3f2](http://feeds.stgsrv.de/index.php?id=26&tx_ttnews[tt_news]=969&cHash=bb42b87b90ba01792e8d29e8e78aa3f2), Abgerufen am 09.06.15.
- [39] Gustafsson, H. and Griebisch, C. (2008). Strategy Analytics: In den Navigationsmarkt kommt Bewegung, In pressebox.de, <http://www.pressebox.de/pressemitteilung/strategy-analytics/Strategy-Analytics-In-den-Navigationsmarkt-kommt-Bewegung/boxid/158180>, Abgerufen am 09.01.16.
- [40] Han, J. and Kamber, M. (2006). Data mining concepts and techniques. Elsevier/Morgan Kaufmann, Amsterdam.
- [41] Hochstädter, A., Zahn, P. and Breuer, K. (2000). Ein universelles Fahrermodell mit den Einsatzbeispielen Verkehrssimulation und Fahrsimulator. In: 9. Aachener Kolloquium Fahrzeug- und Motorentechnik, 4.-6.10.2000, Aachen.
- [42] Holten, R. and Nittel, D. (2010). E-Learning in Hochschule und Weiterbildung: Einsatzchancen und Erfahrungen. In: Erwachsenenbildung und lebensbegleitendes Lernen : EB LBL, Bertelsmann, Bielefeld.

- [43] Homoet, M. (2004). Institut für Kartographie und Geoinformation der Universität Bonn, http://www.google.de/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0CCEQFjAA&url=http%3A%2F%2Fwww.ikg.uni-bonn.de%2Fvorlesungsarchiv%2FVertiefer_Geoinfo_02_03%2FSeminar_WS02_03%2F04-Michael-Homoet-Map-Matching-Techniken.ppt&ei=o1uxVIjMN8TDPPLTgKAO&usg=AFQjCNHzA66FIohflmJYDDDtY7T3nNPB5g&bvm=bv.83339334, d.ZWU, Abgerufen am 09.01.16.
- [44] IT-Times (2011). Garmin kämpft weiter gegen den Abwärtstrend - Smartphones machen dem Unternehmen das Leben schwer, <http://www.it-times.de/news/garmin-kampft-weiter-gegen-den-abwartstrend-smartphones-machen-dem-unternehmen-das-leben-schwer-29683/>, Abgerufen am 09.01.16.
- [45] Jan, O., Horowitz, A. J. and Peng, Z.-R. (2000). Using Global Positioning System Data to Understand variations in Path Choice. In: Transportation Research Record 1725, pp. 37-44.
- [46] Kellerer, H., Pferschy, U. and Pisinger, D. (2004). Knapsack Problems, Springer, Berlin Heidelberg.
- [47] Kessler, C. et al. (2012). EuroFOT Key Results, In euroFOT final event, http://www.eurofot-ip.eu/download/library/deliverables/eurofot_sp120121212v11dld113_final_report.pdf, Abgerufen am 09.01.16.
- [48] Kopf, M. (2005). Was nützt es dem Fahrer, wenn Fahrerinformations- und assistenzsysteme etwas über ihn wissen? In: Fahrerassistenzsysteme mit maschineller Wahrnehmung, pp. 117-139, Springer, Berlin Heidelberg.
- [49] Kuriyagawa, Y. et al. (2002). A Research on Analytical Method of Driver-Vehicle-Environment System for Construction of Intelligent Driver Support System. In: Vehicle System Dynamics: International Journal of Vehicle Mechanics and Mobility, 37(5).
- [50] Landau, A. (2010). Herausforderungen und Chancen für die Datenbasis von intelligenten adaptiven Fahrerassistenzsystemen. In: Beiträge zum Symposium AAET 2010, pp. 363-382, 10.-11.02.10, Braunschweig.
- [51] Landau, A., Ehmke, J. F. and Mattfeld, D. C. (2012). Analyse des Routenwahlverhaltens von Autofahrern mit Geografischen Informationssystemen. In: Tagungsband der Geoinformatik 2012 Mobilität und Umwelt, pp. 91-100, Braunschweig.

- [52] Landau, A., Ehmke, J. F. and Mattfeld, D. C. (2013). Personalization of Travel Times for Personal Vehicle Navigation Systems. In: Proceedings of European Navigation Conference, 23. - 25. April 2013, Wien.
- [53] Landau, A., Ehmke, J. F. and Mattfeld, D. C. (2014a). Bereitstellung von fahrertypabhängigen Kantengewichten zur personalisierten Routenberechnung. In: Tagungsband der MKWI 2014, 1437 - 1485, Paderborn
- [54] Landau, A., Ehmke, J. F. and Mattfeld, D. C. (2014b). Multi-Perspective Data Analysis of Drivers Navigation Behaviour. In: Proceedings ECIS 09.-11-06.2014, Tel Aviv.
- [55] Laporte, G. (1992). The vehicle routing problem: An overview of exact and approximate algorithms. In: European Journal of Operational Research, 59(3), pp. 345-358.
- [56] Li, H. (2014). Investigation Morning Commute Route Choice Behavior Using Global Positioning Systems and Multi-Day Travel Data, Georgia Institute of Technology, Atlanta.
- [57] Little, A. D., (2010). Zukunft der Mobilität 2020 http://www.adlittle.de/uploads/tx_extthoughtleadership/ADL_Zukunft_der_Mobilitaet_2020_Langfassung.pdf, Abgerufen am 09.01.2016.
- [58] Mansfeld, W. (2010). Satellitenortung und Navigation: Grundlagen und Anwendung globaler Satellitennavigation, Vieweg+Teubner, Wiesbaden.
- [59] Marbân, Ô., Mariscal, G. and Segovia, J. (2009). A Data Mining & Knowledge Discovery Process Model. In: Data Mining and Knowledge Discovery in Real Life Applications, I-Tech Education and Publishing, http://www.intechopen.com/books/data_mining_and_knowledge_discovery_in_real_life_applications/a_data_mining_amp_knowledge_discovery_process_model, Abgerufen am 09.01.16.
- [60] Matschke, I. (2007). Einfluss dynamischer Navigation auf das Verkehrsgeschehen in städtischen Straßennetzen, Hannover Technische Informationsbibliothek und Universitätsbibliothek, Hannover.
- [61] Mattsson, M. (2012). Investigating the factorial invariance of the 28-item DBQ across genders and age groups: An Exploratory Structural Equation Modeling Study. In: Accident Analysis & Prevention, Volume 48, pp. 379-396.
- [62] McGinty, L. and Smyth, B. (2001). Collaborative Case-Based Reasoning: Applications in Personalised Route Planning. In: Proceedings of 4th Interna-

- tional Conference on Case-Based Reasoning, ICCBR 2001, pp. 362-376, 30. Juli-2. August 2001, Vancouver, BC.
- [63] Metz, B., Schoch, S. and Krüger, H.-P. (2011). Von der Hypothese zur Auswertung: Planung der Datenauswertung in einem Field Operational Test. In: Tagungsband der 6.VDI-Tagung „Der Fahrer im 21. Jahrhundert“, pp. 317-326, 08.-09.11.11, Braunschweig.
- [64] Metz, B., Landau, A., Hargutt, V. and Neukum, A. (2013). Naturalistic Driving Data - Re-Analyse von Daten aus dem EU-Projekt euroFOT. In: FAT-Schriftenreihe 256: Verband der Automobilindustrie.
- [65] Miyazaki T. et al. (2001). Modeling of Human Behaviors in Real Driving Situations. In: IEEE Intelligent Transportation Systems Conference Proceedings, pp. 643-646, 25.-29. August 2001, Oakland.
- [66] Motortrader editorial team (2013). Market for GPS route navigation systems for cars to grow, <http://www.motortrader.com.my/news/market-for-gps-route-navigation-systems-for-cars-to-grow/>, Abgerufen am 09.01.16.
- [67] Nagaki, K. (2012). Evolution of In-Car Navigation Systems. In: Handbook of Intelligent Vehicles, pp. 463-487, Springer, London.
- [68] Nagel K. and Schreckenberg M. (1992). A cellular automaton model for freeway traffic, In Journal of Physics I France 2, pp. 2221-2229.
- [69] Neumerkel, D. et al. (2002). Fahrermodelle - Ein Schlüssel für unfallfreies Fahren? In: Künstliche Intelligenz - Schwerpunkt: Adaptivität und Benutzermodellierung, 2002-3, p. 34 ff.
- [70] Nowicki, J. (2002). Marktentwicklung für Kfz-Navigation - ADAC-Schätzung und Prognose bis 2005, In ADAC - Verkehrspolitik und Verbraucherschutz, http://www.autohaus.de/fm/2353/adac_prognose.pdf, Abgerufen am 09.01.16.
- [71] Ober-Sundermeier, A. (2003). Entwicklung eines Verfahrens zur Stauprognose an Engpässen auf Autobahnen unter besonderer Berücksichtigung von Arbeitsstellen. In: Schriftenreihe Verkehr, Heft 15, Kassel.
- [72] OpenStreetMap (2015). <http://www.openstreetmap.org/#map=5/51.500/-0.100>, Abgerufen am 09.06.15.
- [73] Pang, G.K.H. (2009). Route Selection in a Neuro-Fuzzy Vehicle Navigation System, In Computational Intelligence in Systems and Control Designs and Applications, p.201-210, Norwell.

- [74] Park, K. et al. (2007). Learning user preferences of route choice behaviour for adaptive route guidance. In: IET Intelligent Transport Systems, 2007, 1(2).
- [75] Plöchel, M. and Edelmann, J. (2007). Driver models in automobile dynamics application, In Vehicle System Dynamics: International Journal of Vehicle Mechanics and Mobility, 45(7-8), pp. 699-741, London.
- [76] Pocketnavigation.de (2008). Garmin weiterhin Marktführer, <http://www.pocketnavigation.de/2009/03/garmin-weiterhin-marktfuehrer/>, Abgerufen am 09.01.16.
- [77] Pocketnavigation.de (2012). Navi-Markt bricht weiter ein, <http://www.pocketnavigation.de/2012/07/navi-markt-bricht-weiter-ein/>, Abgerufen am 09.01.16.
- [78] Quddus, M. A., Ochieng, W. Y. and Noland, R. B. (2007). Current map-matching algorithms for transport applications: State-of-the art and future research directions. In: Transportation Research Part C: Emerging Technologies, 15(5), pp. 312-328.
- [79] Reinhardt, W. (2009). Map Matching. In: Vorlesung Geoinformatik, http://www.unibw.de/inf4/professuren/geoinformatik/lehre/skripten/skripte/skripten.ht_09/map-matching-2009.pdf, Abgerufen am 09.01.16.
- [80] Richter, K.-F., Klippel, A. and Freska, C. (2004). Shortest, Fastest but what next? A Different Approach to Route Directions, Geoinformation und Mobilität - von der Forschung zur praktischen Anwendung. In: Beiträge zu den Münsteraner GI-Tagen, pp. 205-217.
- [81] Runte, M. (2000). Personalisierung im Internet: individualisierte Angebote mit Collaborative Filtering, Deutscher Universitätsverlag, Wiesbaden
- [82] Sandkühler, D. (2004). Verkehrssituationserkennung: Entwicklung und Anwendung in Fahrerassistenzsystemen. In: Beiträge zur VDI-Tagung „Integrierte Sicherheit und Fahrerassistenzsysteme“, 27.-29.10.2004, Wolfsburg.
- [83] Schraut, M. P. (2000). Umgebungserfassung auf Basis lernender digitaler Karten zur vorausschauenden Konditionierung von Fahrerassistenzsystemen, Fakultät für Elektrotechnik und Informationstechnik TU München.
- [84] Schrijver, A. (2004). Combinatorial Optimization - Polyhedra and Efficiency. Algorithms and Combinatorics 24. Springer, Berlin Heidelberg.
- [85] Schulz, A., Fröming, R. and Schindler, V. (2007). Darstellung fahreradaptiver Warn- und Bremsassistentenstrategien durch Online-Analyse des Fahrer-

- verhaltens. In: Beiträge zur VDI-Tagung „Der Fahrer im 21. Jahrhundert“, Braunschweig, 14.-15.11.2007.
- [86] Storandt, S. (2012). Algorithms for Vehicle Navigation, Fakultät Informatik, Elektrotechnik und Informationstechnik Universität Stuttgart.
- [87] Su, X. and Khoshgoftaar, M. (2009). A Survey of Collaborative Filtering Technologies. In: Advances in Artificial Intelligence, Volume 2009.
- [88] Sugiyama, Y. et al (2008). Traffic jams without bottlenecks-experimental evidence for the physical mechanism of the formation of a jam. In: New Journal of Physics.
- [89] Tuan, D. (2012). Recommender systems - How they work and their impacts, <http://findoutyourfavorite.blogspot.de/2012/04/collaborative-filtering.html>, Abgerufen am 09.01.16.
- [90] Tittmann, P. (2003). Graphentheorie: Eine anwendungsorientierte Einführung, Carl Hanser Verlag, München Wien.
- [91] TomTom International BV (2015). Presseinformationen, <http://corporate.tomtom.com/releases.cfm>, Abgerufen am 09.01.16.
- [92] Völkel, T. (2008). Multimodale Annotation geographischer Daten zur personalisierten Fußgängernavigation, <http://nbn-resolving.de/urn:nbn:de:bsz:14-ds-1239804877252-19609>, Abgerufen am 09.01.16.
- [93] Volkmann, L. (1991). Graphen und Digraphen: Eine Einführung in die Graphentheorie, Springer, Wien.
- [94] Vollmuth, J. (2012). Handy-Navigationssysteme - Rund 38 Prozent aller Handys in Europa bis 2012 GPS-fähig, In Elektronik-Praxis, <http://www.elektronikpraxis.vogel.de/marktzahlen/articles/111646/>, Abgerufen am 09.01.16.
- [95] Vollrath, M. (2008). In 500 Metern rechts abbiegen - Wann stören Navigationsinformationen den Fahrer? In: Fortschritte der Verkehrspsychologie - Beiträge vom 45. Kongress der Deutschen Gesellschaft für Psychologie, pp. 259-273, VS Verlag für Sozialwissenschaften, Wiesbaden.
- [96] Wahlberg, A. E., Dorn, L. and Kline, T. (2011). The Manchester Driver Behaviour Questionnaire as a Predictor of Road Traffic Accidents. In: Theoretical Issues in Ergonomics Science, 12(1), pp. 66-86.
- [97] Wang, P. P., Ruan D. and Kerre, E. E. (2010). Fuzzy logic: a spectrum of theoretical & practical issues: With 28 tables, Springer, Berlin Heidelberg.

- [98] Weiß, C. (2006). Datenanalyse und Modellierung mit STATISTICA, Oldenbourg Wissenschaft, München.
- [99] Wendel, J. (2007). Integrierte Navigationssysteme: Sensordatenfusion, GPS und Inertiale Navigation, Oldenbourg Wissenschaft, München.
- [100] Yano Research Institute Ltd. (2011). World Car Navigation & PND Market: Research Findings 2010, <http://www.yanoresearch.com/press/pdf/730.pdf>, Abgerufen am 09.01.16.
- [101] Yao, X. and Fickas, S. (2007). Pedestrian Navigation System: A Case Study for Deep Personalization. In: Proceedings of First International Workshop on Software Engineering. for Pervasive Computing Applications, Systems, and Environments, 20-26 May 2007, Washington.

